International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 4s
DOI: https://doi.org/10.17762/ijritcc.v11i4s.6310
***Article Received**: 24 December 2022 **Revised**: 26 January 2023 **Accepted**: 02 February 2023*

Automatic Detection of Road Cracks using EfficientNet with Residual U-Net-based Segmentation and YOLOv5-based Detection

Satheesh Kumar Gooda[1*], Narender Chinthamu[2], Dr. S. Tamil Selvan[3], Dr. V. Rajakumareswaran[4], Gokila Brindha Paramasivam[5]

[1]Senior Manager, WESCO International, USA, Student of Osmania University, India.
e-mail: skgooda@gmail.com
[2]Enterprise Architect, MIT CTO Candidate
e-mail: Narender.chinthamu@gmail.com
[3]Assistant Professor, Department of Computer Science and Engineering,
Erode Sengunthar Engineering College,
Thudupathi, Perundurai - 638 057, India.
e-mail: Stamilselvancse@gmail.com
[4]Assistant Professor, Department Of Computer Science and Engineering,
Erode Sengunthar Engineering College, Thuduppathi, Tamil Nadu, India,
e-mail: mailtoraja@gmail.com
[5]Assistant Professor Sr.G, Department of Computer Technology UG,
Kongu Engineering College,
Perundurai, Tamil Nadu-638060, India.
Email: brindha.ctug@kongu.edu

Abstract— The main factor affecting road performance is pavement damage. One of the difficulties in maintaining roads is pavement cracking. Credible and reliable inspection of heritage structural health relies heavily on crack detection on road surfaces. To achieve intelligent operation and maintenance, intelligent crack detection is essential to traffic safety. The detection of road pavement cracks using computer vision has gained popularity in recent years. Recent technological breakthroughs in general deep learning algorithms have resulted in improved results in the discipline of crack detection. In this paper, two techniques for object identification and segmentation are proposed. The EfficientNet with residual U-Net technique is suggested for segmentation, while the YOLO v5 algorithm is offered for crack detection. To correctly separate the pavement cracks, a crack segmentation network is used. Road crack identification and segmentation accuracy were enhanced by optimising the model's hyperparameters and increasing the feature extraction structure. The suggested algorithm's performance is compared to state-of-the-art algorithms. The suggested work achieves 99.35% accuracy.

Keywords- Crack, Segmentation, Object detection, Deep learning, EfficientNet, U-Net, YOLO v5.

I. INTRODUCTION

Detecting crack damage in industrial and civil constructions has long been an issue. Manual and machine detection methods are used in traditional crack detecting technologies. The consensus is that manual detection takes longer and is less reliable. In recent years, the technology for machine detection methods based on ultrasonic, microwaves, or other signals has improved rapidly [1-3]. Due to these restrictions, both commercial and academic institutions have been researching autonomous crack detection techniques [4]. Due to the widespread availability of smartphones and cameras, image-based approaches are deemed to be greatly cost-effective [5].

The effectiveness of computer vision methods has been demonstrated in automating the image-based crack detection approach, and their use has become a research problem in recent decades [6]. As a result, image-based crack localization investigations are broadly classified as manual or automatic feature extraction-based methodologies. A computer vision method for crack detection begins by identifying crack sensitive features, which can be accomplished using deep learning approaches or image processing techniques (IPTs). The employment of techniques like edge detectors, morphological processes, and thresholding was the subject of early study.

In addition to noise, different illumination conditions can affect the techniques for manually identifying cracks based on

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 4s
DOI: https://doi.org/10.17762/ijritcc.v11i4s.6310
***Article Received**: 24 December 2022 **Revised**: 26 January 2023 **Accepted**: 02 February 2023*

feature extraction. Due to this problem, deep architectures are now used for crack detection problems that don't require custom characteristics [7]. A deep architecture uses several deep layers to extract high-level properties from raw inputs, making it the "next generation" of neural networks. Objects of interest in images may be found by using bounding boxes in computer vision. The sliding window method was found to be less accurate when it came to localizing and recognizing items in images.

In the crack detection section, bounding boxes are used to detect crack areas on input photographs [9]. Computer vision approaches used during implementation of OR include region-based convolutional neural networks (R-CNNs) [9], you only look once (YOLO) [10], and single shot detectors (SSDs) [11]. Crack detection frequently uses members of the R-CNN family. SSD and YOLO designs have only been used once as primary frameworks for crack detection, according to the authors. Identifying cracks in photographs may be viewed as an object detection and classification issue. As a consequence, a deep learning-based model might be used to detect surface crack faults in pavement and bridges. To build up an automatic crack detection system, four steps must be completed: image collection, image pre-processing, image segmentation, and crack detection. The contribution of this research is,

- For object detection, YOLOv5 model is used in this work.
- An EfficientNet with residual U-Net is used to segment the cracks in the road images.

The organization of the work is as follows. Section 2 presents the related works, Section 3 describes the methodology, Section 4 discusses the results and Section 5 concludes the work.

II. RELATED WORKS

Numerous studies have concentrated on using a multi-class classification strategy to solve the issue, taking both crack identification and crack type categorization into consideration [12–14]. According to Park et al. [14], this multi-class classification technique based on CNN was used to categorize road images into crack, intact regions and road markers. The study categorized crack types into five types which were influenced by the AlexNet and LeNet networks as well as assessed and compared four CNNs with varied depths. Using deep architectures and handcrafted features as the main criteria for crack investigation, [15,16] compared crack investigate techniques. Deep architectures were compared to the effectiveness of a number of edge detectors, including Sobel, Canny, Prewitt, Butterworth, and others.

A comparison between Hessian matrix and Haar wavelet accelerated robust features technique and convolutional neural network extraction for crack recognition was conducted by Kim et al. [17]. In [18], features were learned using pretrained VGG-16 and AlexNet models on the ImageNet dataset. As an alternative to conventional machine learning approaches, we used fully connected layers and soft-max layers for classifying the attributes. According to Li et al. [19], a multiscale defect region proposal network (RPN) generates candidate bounding boxes at various levels to increase detection accuracy. The authors used a second deep architecture and a geotagged picture database to perform geolocalization. It is pertinent to note that the geo-localization module and the crack detection network are part of the same network.

An improved version of the fast R-CNN has been proposed for crack detection. To speed up training, a CNN is combined with a sensitivity detection network to extract deep features. On the other hand, Deng et al. [21] evaluated the dataset of concrete that included pictures with handwritten typescripts. They concluded that handwritten characters may be considered excessive noise in solid images. An OR setup was used by Maeda et al. [22] to find cracks in a big data set built using an SDD architecture. MobileNet and Inception V2 form the backbone of the SSD framework's feature extraction. It is important to emphasise that the data set was obtained and annotated by researchers.

The algorithms used by Ni et al. [23] o identify crack repairs were GoogleNet and ResNet. To partition the observed crack locations, Otsu's thresholding was utilised, accompanied by median filtering and Hessian matrices. These methods are performed to reduce the impact of brightness and to enhance crack structures. [24] used transfer learning with a pre-trained framework on the ImageNet data set to identify crack patches. Quick blockwise segmentation and vector voting curve detection techniques were then used to produce the crack mask and increase crack localization accuracy. GoogLeNet was used in [25] to anticipate crack fixes. To segment the cracks, the discovered patches were passed through a feature fusion component and a number of convolution layers. Zhang et al. [26] proposed a computationally more efficient Sobel-edge adaptive sliding window strategy for obtaining crack patch than the traditional sliding window method.

III. METHODOLOGY

A crack detecting approach based on a deep learning approach is proposed in this research. The image database is first organised, and image noise in the dataset is filtered away to improve the contrast between road cracks and backgrounds. Following that, the filtered photos are sent into a crack prediction model for training. Figure 1 depicts the flow of the proposed work.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 4s
DOI: https://doi.org/10.17762/ijritcc.v11i4s.6310
***Article Received**: 24 December 2022 **Revised**: 26 January 2023 **Accepted**: 02 February 2023*

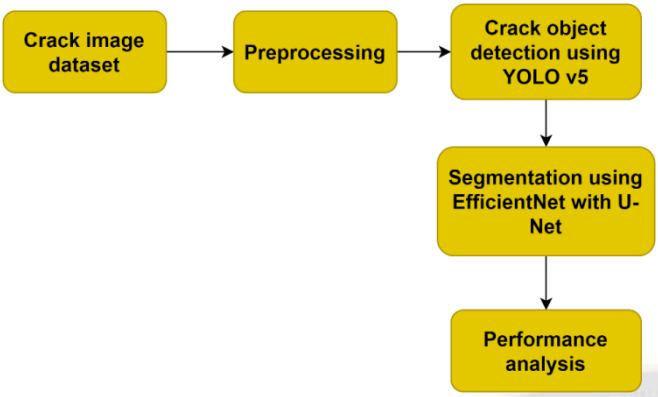

Figure 1. Flow of the proposed road crack detection

3.1 CFD Data Set

The CFD dataset was gathered and made available to the public [27]. The data collection, according to the authors, usually represents the status of the urban road surface in Beijing, China. The photos have been carefully labelled down to the pixel level. The CFD comprises 118, 480 × 320 pixel RGB road photos with noise like shadows, water stains, and oil patches as well as various lighting circumstances. The images were captured with an iPhone 5 with a focus of 4mm, an exposure of f/2.4, and an exposure duration of 1/134 s. The image's width, which varies from 1 to 3 mm, should be noted. Figure. 2 displays a few photos of road cracks.

Figure 2. Road crack images

3.2 *Image preprocessing*

Three-channel colour photographs of road cracks were used. Red, green, and blue made up each of the three colours that make up a colour pixel. These three colours each had a connection to a colour image at a particular spatial location, which led to the creation of a vector to depict the image. Two colour augmentation techniques (contrast and sharpness) were utilised to process colour photographs. Contrast augmentation was employed to expand the gray-level range and enhance image clarity to address the low contrast issue brought on by the crack picture's constrained gray-level range. The probability approach was used to smooth the intensity and saturation components of the hue-saturation-intensity (HSI) colour model, resulting in a uniform distribution. Eqs. (1) and (2) illustrate the computation procedures for the brightness and saturation components, respectively.

$$y_{1k} = F(x_{1k}) = P\{x_I \le x_{IK}\} = \sum_{m=0}^{k} f(x_{Im}) = \sum_{m=0}^{k} P\{x_I \le x_{IK}\} \quad (1)$$

$$y_{st} = F(x_{1k}|x_{IK}) = \sum_{m=0}^{t} f(x_{Sm}|x_{IK}) = \sum_{m=0}^{t} \frac{P\{x_I = x_{IK}, x_S = x_{Sm}\}}{P\{x_I = x_{Ik}\}} \quad (2)$$

where k = 0, 1, . . . , L-1 and t = 0, 1, . . . , M-1; L and M denote discrete levels of intensity and saturation, respectively. $\mathbf{X} = (\mathbf{x}_H, \mathbf{x}_S, \mathbf{x}_I)T$ is a vector of color pixels representing each image. F(.) is a probability function, and $F(\mathbf{Z}) = F(x_I, x_S) = P\{\mathbf{x}_I \le x_I, \mathbf{x}_S \le x_S\}$.

By boosting the contrast of the surrounding pixels, sharpness reduces the blurring of the picture's object and defines it. For the crack image, Laplace sharpness produces gradient values (Laplace operator). Eq.(3) illustrates the enhancement technique based on the Laplace operator.

$$p(a,b) = q(a,b) - \begin{bmatrix} \nabla^2 R(a,b) \\ \nabla^2 G(a,b) \\ \nabla^2 B(a,b) \end{bmatrix} \quad (3)$$

where p (a,b) represents the sharpened crack picture and q (a,b) depicts the original crack picture, and $\nabla^2 R(a,b)$, $\nabla^2 G(a,b)$, and $\nabla^2 B(a,b)$ are the red, green, and blue components' respective Laplace operators in colour pictures.

3.3 *Segmentation using EfficientNet with residual U-Net Architecture*

A proposed Efficient-U-Net network consists of an encoder and decoder, as illustrated in Figure 4. Due to limited resources, we utilize a modified EfficientNetB4 encoder. There are 9 stages in the encoder: 3x3 convolutional layers, 32 mobile reverse bottleneck convolutional structures, and 11 convolutional layers. There are five upsampling processes and a sequence of convolutions in the decoder. In order to determine the segmentation results, the encoder restores the original picture size based on the retrieved features. We limit the noise response and concentrate on specific properties of the segmented crack by adding an attention gate to the skip connection. The network may be expanded by including the residual structure. After each convolution, the residual block applies batch normalization (BN) and ReLU activation. Batch normalization reduces gradient propagation and vanishment and accelerates network convergence. Non-linear processing using ReLU can be used to expand the network's capability to express itself non-linearly. Figure 3 displays the segmented image.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 4s
DOI: https://doi.org/10.17762/ijritcc.v11i4s.6310
***Article Received**: 24 December 2022 **Revised**: 26 January 2023 **Accepted**: 02 February 2023*

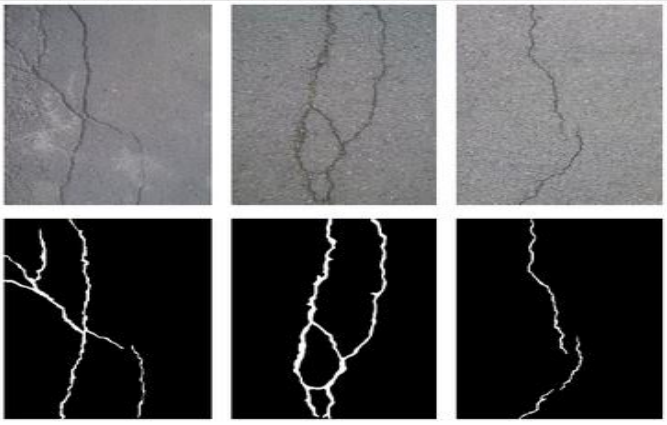

Figure 3. Original and segmented images

The dropout layer, 1x1 convolutions for dimension reduction, SE modules, depthwise convolutions, and 1x1 convolutions are all incorporated in the MBConv structure. BN and Swish activation processes are carried out after the first 1x1 convolution and Depthwise convolution, whereas BN operations are only carried out in the second 1x1 convolution. A shortcut link is combined with additional feature information. The shortcut link will appear if the output and input MBConv structure feature matrices are identical. The accuracy of target recognition, picture segmentation, and image classification have all been greatly enhanced by the SE module. A Sigmoid activation function, a global average pooling, and two fully connected layers were all used in this investigation. Swish activation is introduced between two layers that are completely connected. Stretch an image with HWC compression into a 1x1xC format utilising global pooling and fully connected layers, then multiply the resulting image by the input image to give each channel weight. In this method, the SE module allows the network to learn more about crack-related features. An attention gate is a type of attention device that may automatically focus on a certain area, muffle the response of unnecessary regions, and enhance feature data that is essential to a certain task.

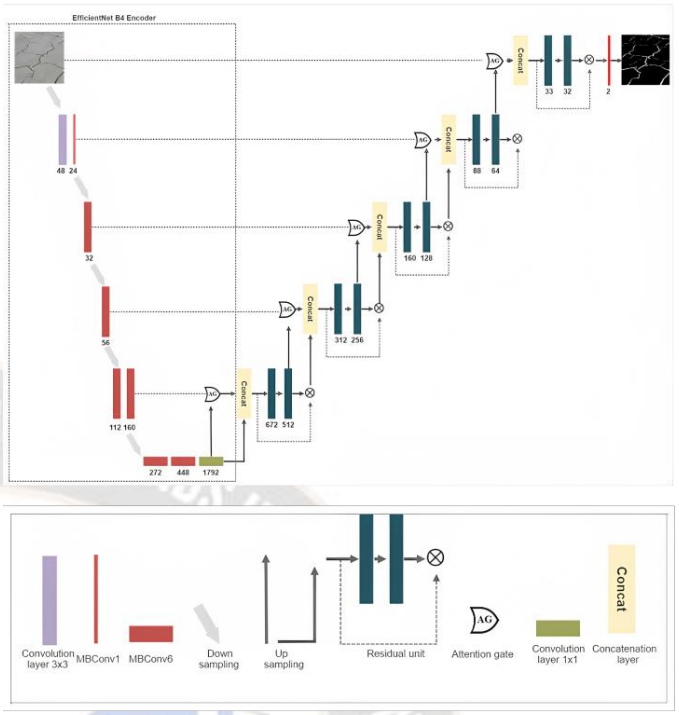

Figure 4. The architecture of the proposed EfficientNet-U-Net

3.4 *Object detection using YOLOv5*

There are four versions of YOLOv5, each with varying detection methods: YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x. With a weight of 13.7 M, and a parameter of 7.0 M, YOLOv5s is the fastest and smallest model. A framework of the algorithm is illustrated in Figure 5, which consists of three parts: the bottleneck, the backbone, and detection part. Four modules make up the backbone network: the focus module (Focus), the standard convolution module (Conv), the C3 module, and the spatial pyramid pooling module (SPP). As part of the YOLOv5 network structure adjustment, two parameters are adjusted: depth factor and width factor. Because it is a one-stage network with multilayer feature map prediction, the YOLOv5s approach provides great accuracy and detection speed. It may be utilised effectively in industry and satisfies the criteria for pavement crack detecting operations, especially in terms of speed. YOLOv5s model performance is improved, model size is reduced, and detection accuracy is increased with a lightweight network structure based on accuracy, parameter number, and computational cost.

A convolutional neural network called Backbone is created by combining visual input with different particle sizes. Head processes box and class prediction procedures by incorporating features from Neck (PAnet) and Feature Pyramid Network (FPN). Neck is a layer sequence that combines and integrates picture characteristics to provide prediction. The FPN structure improves detection of multi-scale items while providing an effective trade-off between identification speed and accuracy. Focus and Cross-

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 4s
DOI: https://doi.org/10.17762/ijritcc.v11i4s.6310
***Article Received**: 24 December 2022 **Revised**: 26 January 2023 **Accepted**: 02 February 2023*

Stage Partial Connections (CSP) are the most important features of YOLOv5. The focus layer was developed to improve forward and backward performance, lessen the effect of mAP, and minimize the number of layers, parameters, FLOPS, and CUDA memory.

The latest YOLOv5 version and its predecessor have two key changes. First, swap out the Focus layer for a 6 x 6 Conv2d layer. It has the same properties as a conventional 2D convolution layer, despite lacking space-to-depth functionality. Convolution layers with kernel sizes of six and stride two. Focus layers with kernel sizes of three. The second change was to replace the SPP layer with the SPPF layer. These actions nearly triple computer performance. This alternative is therefore quicker and more effective. The main layer of the original YOLOv5 structure, the Conv layer, was analyzed and changed. In the first Conv layer, an activation function called SiLU (Sigmoid-Weighted Linear Units) was used.

The Conv layer frequently employs ReLU as an activation function (Rectified Linear Unit). Due to the minimal processing required, learning occurs quickly, and implementation is straightforward. The ReLU activation function has the drawback that if it produces a value less than zero, the gradient and weight will presumably remain at zero during learning. As a consequence, we altered a structure of the Conv layer. As a result, there is also the negative aspect of ineffective learning. A variation of the ReLU activation function is the ELU activation function. This shortens training time and improves the performance of neural network test datasets.

$$ELU(x) = \begin{cases} x & if\ x>0 \\ \alpha \exp(x) - 1, & if\ x\leq 0 \end{cases} \quad (4)$$

$$ELU'(x) = \begin{cases} 1 & if\ x>0 \\ f(x) + \alpha, & if\ x\leq 0 \end{cases} \quad (5)$$

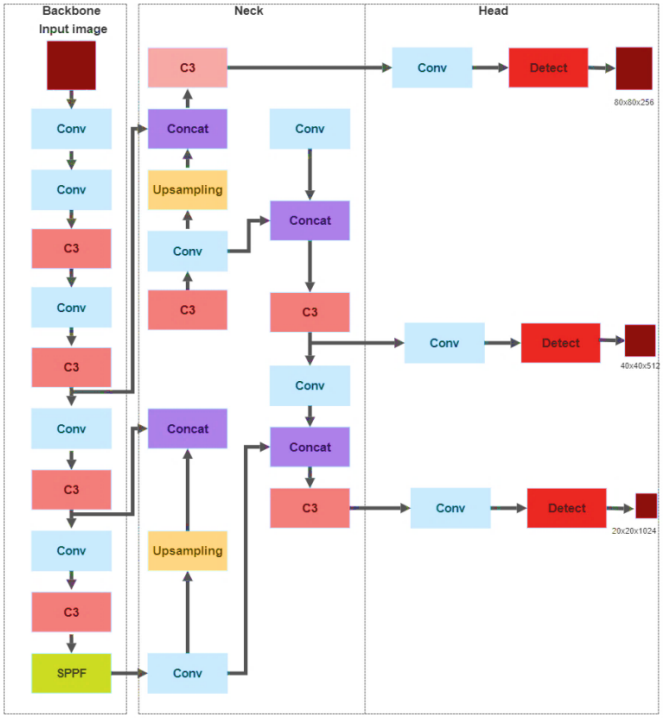

Figure 5. Structure of YOLOv5 network

IV. EXPERIMENTAL RESULT AND DISCUSSION

Simulation tests were carried out on the CFD dataset using an NVIDIA TESLA P100 GPU and 16 GB RAM in this study. The suggested method was created using Pytorch and Tensorflow in a Python environment on a Linux platform. To test the suggested models, five performance measures were generated using Equations (6) to (11), namely Jaccard coefficient, Dice coincidence index (Sorensen similarity coefficient), accuracy, precision, recall, and IoU.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (6)$$

$$Precision = \frac{TP}{TP+FP} \quad (7)$$

$$Recall = \frac{TP}{TP+FN} \quad (8)$$

$$Dice_{index} = \frac{2\times Precision\times Recall}{Precision+Recall} \quad (9)$$

$$IoU = \frac{GroundTruth\cap Prediction}{GroundTruth\cup Prediction} \quad (10)$$

$$Jaccard = \frac{TP}{TP+FP+FN} \quad (11)$$

To compare the performance several state-of-the-art algorithms are used to evaluate the proposed work. All models were trained for a total of 100 epochs. For the CrackNet, Deep Crack, Deep ResU-Net, and ResU-Net++ models, as well as the proposed Efficient-U-Net model, training began with a batch size of 32.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 4s
DOI: https://doi.org/10.17762/ijritcc.v11i4s.6310
***Article Received**: 24 December 2022 **Revised**: 26 January 2023 **Accepted**: 02 February 2023*

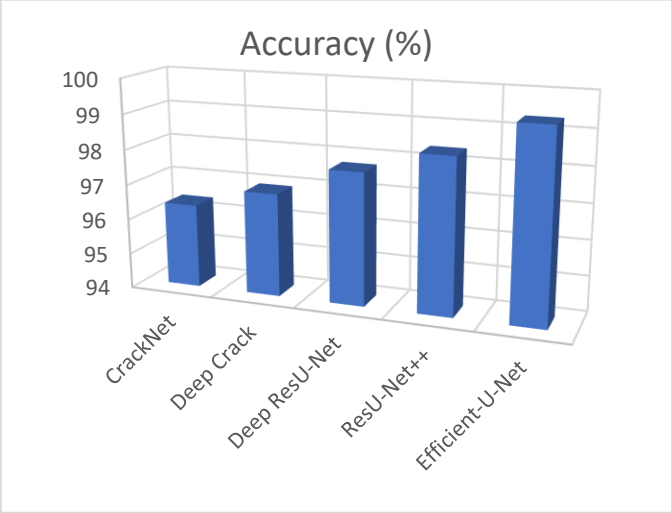

Figure 6. Accuracy comparison of the algorithms

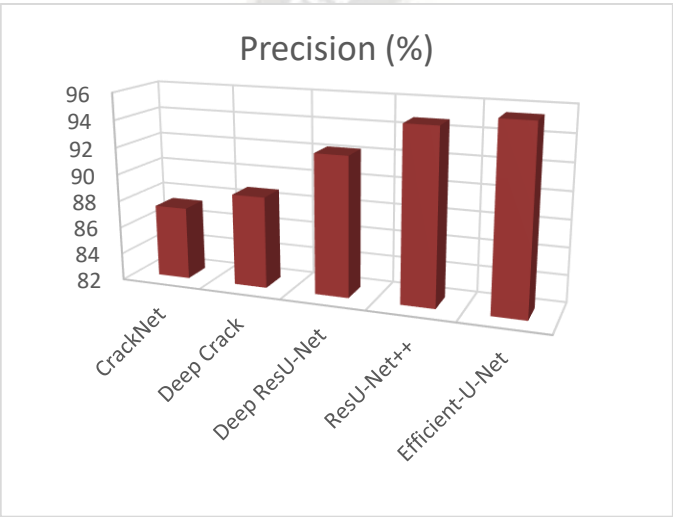

Figure 7. Precision comparison of the algorithms

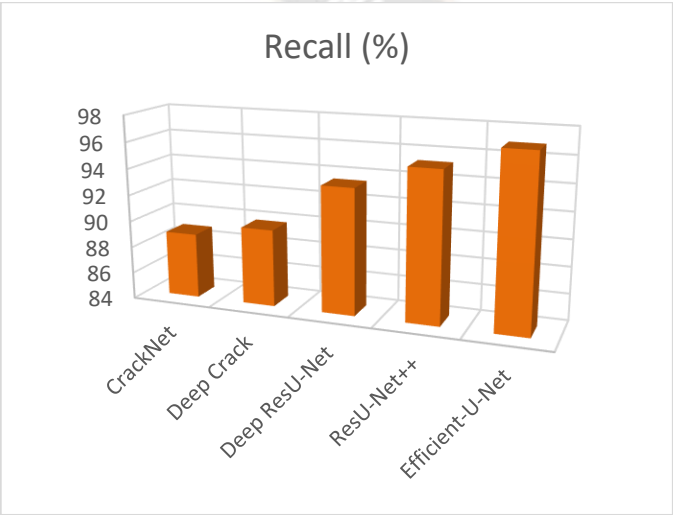

Figure 8. Recall comparison of the algorithms

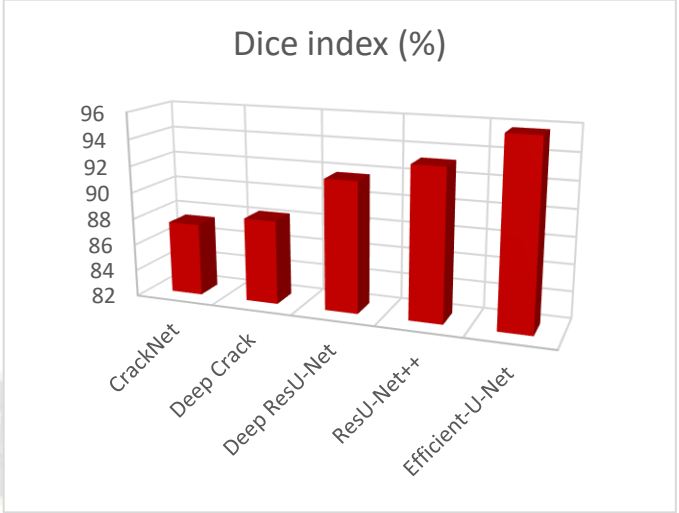

Figure 9. Dice score comparison of the algorithms

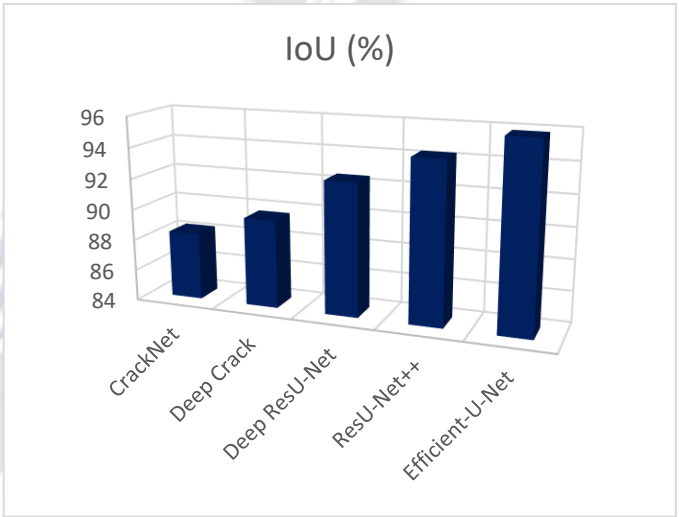

Figure 10. IoU score comparison of the algorithms

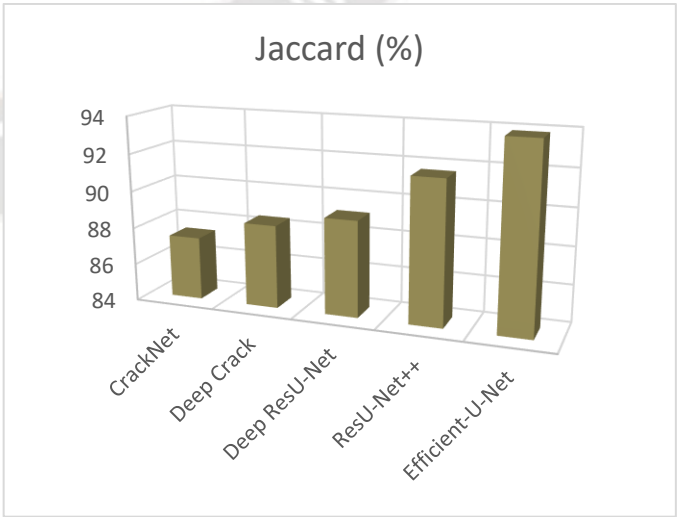

Figure 11. Jaccard score comparison of the algorithms

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 4s
DOI: https://doi.org/10.17762/ijritcc.v11i4s.6310
***Article Received**: 24 December 2022 **Revised**: 26 January 2023 **Accepted**: 02 February 2023*

The accuracy of the algorithms are compared as shown in Figure. 6, for analyzing the performance of the proposed system in crack detection. The accuracy obtained by the proposed Efficient-U-Net is 99.35% which is 0.97% higher than ResU-Net++, 1.6% higher than Deep ResU-Net, 2.42% higher than Deep Crack, and 2.97% higher than CrackNet. The ResU-Net++ obtained the accuracy of 98.38%, Deep ResU-Net obtains 97.75%, Deep Crack obtains 96.93% and CrackNet obtains 96.38%. The residual unit improves performance because feature accumulation with recurrent residual convolutional layers offers improved feature representation for segmentation tasks. It enables the creation of a superior U-Net architecture with the same amount of network parameters and improved picture segmentation performance. The precision of the algorithms are presented in the Figure.7. It is observed that the proposed Efficient-U-Net obtained the maximum precision of 95.47% and CrackNet, Deep Crack, Deep ResU-Net, and ResU-Net++ obtained 87.37%, 88.78%, 92.28%, and 94.76% respectively. Figure. 8 shows the recall obtained by the different algorithms. The recall score obtained by CrackNet, Deep Crack, Deep ResU-Net, ResU-Net++, and Efficient-U-Net is 88.91%, 89.83%, 93.49%, 95.28%, and 96.97% respectively.

However, it is shown that preprocessing enhances the efficiency of the suggested approach while dealing with blurred pictures, raising the Efficient-U-Net's Dice score to 96.97%. The dice score of other algorithms are 87.47%, 88.32%, 91.78%, 93.26% by CrackNet, Deep Crack, Deep ResU-Net, and ResU-Net++ respectively. The capacity to identify small cracks and unlabeled cracks is the key benefit of utilising the Efficient-U-Net model for crack segmentation. Additionally, the Efficient-U-Net model outperforms other models in terms of detecting cracks in images that are blurry and cracks on edges. The model also has the ability to recognise shadow-producing picture cracks caused by shifting lighting conditions. The IoU score of CrackNet, Deep Crack, Deep ResU-Net, ResU-Net++, and Efficient-U-Net is 88.34%, 89.71%, 92.53%, 94.32%, and 95.87% respectively. The Jaccard score obtained by the CrackNet, Deep Crack, Deep ResU-Net, ResU-Net++, and Efficient-U-Net is 87.39%, 88.48%, 89.16%, 91.67%, and 93.88% respectively. The investigation shows that the suggested approach operates more effectively and has increased efficiency across the board. The segmentation outcomes have been greatly enhanced by the usage of two concatenated encoder-decoder designs.

V. CONCLUSION

The need for intelligent monitoring technology is growing as a result of the rapidly rising road mileage, which makes it impossible for the conventional road crack monitoring approach to keep up with demand. In this work, segmentation and object detection are used to evaluate deep learning techniques for crack detection on roads. This research presents a segmentation network that can identify road cracks using an EfficientNet with residual attention based U-Net architecture. The Efficient-U-Net network is used as the segmentation model and the YOLO v5 network is utilised as the detection model to identify cracks accurately while simultaneously segmenting the cracks in the roads. The issue of inaccurate crack localization in the road crack detection network is resolved by combining the segmentation model with the detection model. Results from the experiments show that the suggested model not only performs better than the other models, but also achieves superior accuracy, precision, and recall. The suggested system's accuracy is 99.35%, which is more than that of any existing methods.

REFERENCES

[1] Lacidogna, G.; Piana, G.; Accornero, F.; Carpinteri, A. Multi-technique damage monitoring of concrete beams: Acoustic Emission, Digital Image Correlation, Dynamic Identification. Constr. Build. Mater. 2020, 242, 118114

[2] Zhao, S.; Sun, L.; Gao, J.; Wang, J. Uniaxial ACFM detection system for metal crack size estimation using magnetic signature waveform analysis. Measurement 2020, 164, 108090.

[3] Zhang, X.; Wang, K.; Wang, Y.; Shen, Y.; Hu, H. Rail crack detection using acoustic emission technique by joint optimization noise clustering and time window feature detection. Appl. Acoust. **2020**, 160, 107141.

[4] Cha, Y.J.; Choi, W.; Suh, G.; Mahmoudkhani, S.; Büyüköztürk, O. Autonomous Structural Visual Inspection Using Region-Based Deep Learning for Detecting Multiple Damage Types. Comput.-Aided Civ. Infrastruct. Eng. 2017, 33, 731–747.

[5] Fang, F.; Li, L.; Gu, Y.; Zhu, H.; Lim, J.H. A novel hybrid approach for crack detection. Pattern Recognit. **2020**, 107, 107474.

[6] Yang, F.; Zhang, L.; Yu, S.; Prokhorov, D.; Mei, X.; Ling, H. Feature Pyramid and Hierarchical Boosting Network for Pavement Crack Detection. IEEE Trans. Intell. Transp. Syst. 2020, 21, 1525–1535.

[7] Fang, F.; Li, L.; Gu, Y.; Zhu, H.; Lim, J.H. A novel hybrid approach for crack detection. Pattern Recognit. 2020, 107, 107474.

[8] Hsieh, Y.A.; Tsai, Y.J. Machine Learning for Crack Detection: Review and Model Performance Comparison. J. Comput. Civ. Eng. 2020, 34, 04020038.

[9] Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. arXiv 2013, arXiv:1311.2524.

[10] Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. arXiv 2016, arXiv:1506.02640.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 4s
DOI: https://doi.org/10.17762/ijritcc.v11i4s.6310
***Article Received**: 24 December 2022 **Revised**: 26 January 2023 **Accepted**: 02 February 2023*

[11] Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. SSD: Single Shot MultiBox Detector. arXiv 2016, arXiv:1512.02325.

[12] Li, B.;Wang, K.C.P.; Zhang, A.; Yang, E.;Wang, G. Automatic classification of pavement crack using deep convolutional neural network. Int. J. Pavement Eng. 2020, 21, 457–463.

[13] Feng, C.; Liu, M.Y.; Kao, C.C.; Lee, T.Y. Deep Active Learning for Civil Infrastructure Defect Detection and Classification. Comput. Civ. Eng. 2017, 2017, 298–306.

[14] Park, S.; Bang, S.; Kim, H.; Kim, H. Patch-Based Crack Detection in Black Box Images Using Convolutional Neural Networks. J. Comput. Civ. Eng. 2019, 33, 04019017.

[15] Nhat-Duc, H.; Nguyen, Q.L.; Tran, V.D. Automatic recognition of asphalt pavement cracks using metaheuristic optimized edge detection algorithms and convolution neural network. Autom. Constr. 2018, 94, 203–213.

[16] Dorafshan, S.; Thomas, R.J.; Maguire, M. Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete. Constr. Build. Mater. 2018, 186, 1031–1045.

[17] Kim, H.; Ahn, E.; Shin, M.; Sim, S.H. Crack and Non crack Classification from Concrete Surface Images Using Machine Learning. Struct. Health Monit. 2019, 18, 725–738.

[18] Kim, B.; Cho, S. Automated Vision-Based Detection of Cracks on Concrete Surfaces Using a Deep Learning Technique. Sensors 2018, 18, 3452.

[19] Li, R.; Yuan, Y.; Zhang, W.; Yuan, Y. Unified Vision-Based Methodology for Simultaneous Concrete Defect Detection and Geolocalization. Comput.-Aided Civ. Infrastruct. Eng. 2018, 33, 527–544.

[20] Huyan, J.; Li,W.; Tighe, S.; Zhai, J.; Xu, Z.; Chen, Y. Detection of sealed and unsealed cracks with complex backgrounds using deep convolutional neural network. Autom. Constr. 2019, 107, 102946.

[21] Deng, J.; Lu, Y.; Lee, V.C.S. Concrete crack detection with handwriting script interferences using faster region-based convolutional neural network. Comput.-Aided Civ. Infrastruct. Eng. 2020, 35, 373–388.

[22] Maeda, H.; Sekimoto, Y.; Seto, T.; Kashiyama, T.; Omata, H. Road Damage Detection and Classification Using Deep Neural Networks with Smartphone Images. Comput.-Aided Civ. Infrastruct. Eng. 2018, 33, 1127–1141.

[23] Ni, F.; Zhang, J.; Chen, Z. Zernike-moment measurement of thin-crack width in images enabled by dual-scale deep learning. Comput.-Aided Civ. Infrastruct. Eng. 2019, 34, 367–384.

[24] Zhang, K.; Cheng, H.D.; Zhang, B. Unified Approach to Pavement Crack and Sealed Crack Detection Using Preclassification Based on Transfer Learning. J. Comput. Civ. Eng. 2018, 32, 04018001.

[25] Ni, F.; Zhang, J.; Chen, Z. Pixel-level crack delineation in images with convolutional feature fusion. Struct. Control Health Monit. 2019, 26, e2286.

[26] Zhang, X.; Rajan, D.; Story, B. Concrete crack detection using context-aware deep semantic segmentation network. Comput.-Aided Civ. Infrastruct. Eng. 2019, 34, 951–971.

[27] Shi, Y.; Cui, L.; Qi, Z.; Meng, F.; Chen, Z. Automatic Road Crack Detection Using Random Structured Forests. IEEE Trans. Intell.Transp. Syst. 2016, 17, 3434–3445.