TRANSACTIONS ON ELECTROMAGNETIC SPECTRUM

Defect Segmentation of Magnetic Tiles with the Novel Ardise-U-Net

Çağrı Suiçmez[1*]

[1]Electrical and Electronics Engineering, Faculty of Technology, Gazi University, Ankara, TÜRKİYE

*** Corresponding author's e-mail address:** cagrisuicmez@gmail.com

Received: 22 October 2024
Revised: 28 December 2024
Accepted: 10 January 2025

Research Article
Vol. 4 / No. 1 / 2025
Doi: 10.5281/zenodo.14634587

Abstract: Detecting and diagnosing surface defects is critical throughout the production process. The detection of these defects with high precision plays a major role in preventing both financial and temporal losses. Recently, image processing, machine learning, and deep learning-based approaches have made great progress in surface defect detection. Therefore, in this paper, a new approach called Attention Residual Dilation Squeeze and Excitation U-net (ARdiSE-U-net) is proposed for surface defect detection. The suggested method includes an encoder and decoder, which are the key components of the Unet structure, a depth-wise compression and stimulation block summed to the Unet's skip connections, and a hybridization of attention residual blocks. A model created by applying multiple improvements on the classical U-Net architecture is described. The model integrates attention mechanisms, residual blocks, dilation applications, and Compression-Evoked (SE) blocks. This combination aimed to improve the segmentation performance and provide deeper feature extraction. In this way, critical information is extracted. Lastly, pixel-stage flaw detection is designed with the help of the sigmoid function in the output layer. In the developed model, MT dataset is used for magnetic surface flaw detection. In the experimental studies, Accuracy, Mean IoU, AUC, MAE results of the developed ARdiSE-U-net architecture are 0.9893, 0.8128, 0.8936, 0.0099, respectively. It has higher performance metrics than most of the baseline studies.

Keywords: U-net, Deep learning, Magnetic tiles

Cite this paper as: Suiçmez, Ç., Defect Segmentation of Magnetic Tiles with the Novel Ardise-U-Net. Transactions on Electromagnetic Spectrum. 2025; 4(1): 7-21, Doi:10.5281/zenodo.14634587

1. INTRODUCTION

The detection of surface defects (DSD) is vital in the manufacture of materials and machine components. This is important for both manufacturers and customers. Products free from surface defects will increase the trust of customers and manufacturers towards each other and, at the same time, increase standards. For this reason, removing surface defects in industrial areas such as steel, textiles and marble are very important due to customer satisfaction and increased manufacturer reliability. However, today surface defects are detected manually by humans. Detection performance can be limited due to negative effects such as fatigue and lack of attention. For this reason, unmanned DSD has recently become a popular topic in the literature. At these stages, various image processing techniques are used to reduce the negative effects on humans.

Researchers apply image processing techniques such as using the grayscale value and slope edge, wavelet [1], curvelet [2], and shearlet [3, 4] transformations. At the same time, classic machine learning and image

processing approaches including thresholding [5], gradient [6], morphological filters [7, 8], grey level co-occurrence matrix [9], Fourier [10], and Gabor [11] have been employed to identify surface defects. Deep learning has gained popularity in surface defect identification in recent years and has demonstrated encouraging results in classification challenges [12-14].

This success is due to the ability of deep learning algorithms to automatically extract effective features through convolutional and nonlinear activation layers [1, 15, 16]. In automatic damage detection for magnetic tiles, patterns are often presented which, due to the diversity of defect shape, texture complexity, and illumination conditions, become noises in greyscale images. Since there is a lot of randomness in these patterns, image detection tasks may encounter difficulties. A lot of deep learning models are developed for the detection and prediction of these error patterns. In the first proposed methods, surface defects are presented as present or absent [3, 17]. These proposed methods have achieved high success, but they have not been able to detect and identify the location of surface defects. Some studies have tried to classify the defective parts in the mentioned defects [15, 16]. Nevertheless, these approaches are extremely inefficient since they employ separate CNN models for every defect class. In the literature, the models with the most successful results in both defect classification and localization are segmentation-based fully convolutional network (FCN) models [1, 18]. As opposed to the classical CNN models, FCN network models do not have fully connected layers. In these models, a convolution layer acts as the output stage. In this way, pixel-based segmentation is achieved. In particular, FCN models with skipping connections such as Unet [18] and FPN [19] architectures, are used in the literature [1, 20, 21]. Thanks to the skip-connection properties of these architectures, it is possible to use spatially detailed and difficult-to-obtain features such as shapes, textures, and colors in defect detection. Thanks to these features, segmentation architectures have achieved high success in defect detection.

Among the segmentation algorithms, U-net features are often preferred. The U-net [18] model includes two parts: encoder and decoder. The encoder block includes the convolution and pooling layers of a classical CNN structure. The decoder block contains an upsampling layer instead of a pooling layer. In this way, the feature maps processed by decreasing their dimensions in the encoder block are processed by increasing their dimensions. In U-net models, skip connections merge low-level characteristics covering spatial details in the initial layers with high-level characteristics covering semantic details in the final stages. Thanks to these processes, U-net models have begun to be used frequently for the DSD where location information and pixel diversity are very important [22, 23]. Huang et al. [23, 24], proposed the U-net based MCuePushU network architecture for magnetic tiling DSD. Junfeng et al. [24] combined MobileNetV2 network with U-net and applied it to fabric DSD. Damacharla et al. [22] merged Denesenet and ResNet with the U-net model employing an equivalent method. Luo et al. [25] modified the pre-trained VGG network and LSTM architecture to the Unet system and achieved successful results. One of the main reasons why U-net based model architectures are good at segmentation is that they directly combine low and high-level features.

However, this fusion process also has disadvantages, for example, although low-level feature maps are good for DSD, they may suppress high-level features. In some studies in the literature, some methods have been used to avoid the above problem. One of these methods is to use compression-excitation models [26]. In this way, meaningful information is emphasized. In other studies in the literature, attention gates are used [20, 27-29]. In the last referenced studies, attention blocks and different features have been tried to be combined effectively. However, as mentioned above, in order to prevent the loss of low-level features, low-level features are weighted with a gating mechanism. This process uses a depth separable convolution [30], based compression-excitation structure [31, 32].

In our study, successful results were obtained with the ARdiSE-U-net architecture we proposed for magnetic surface defect detection. The devoloped model was tested on the MT dataset. In the experimental results, ARdiSE-U-net achieves better results than the average of the basic studies. In this study, the contributions of the ARdiSE-U-net model developed for magnetic surface defect detection are as follows:

- The compression-excitation block (SE) is hybridized into the main structure to detect surface defects and to improve performance.

- ARdiSE-U-net architecture is proposed for surface defect detection, and it is introduced for the first time in this study.

- The proposed ARdiSE-U-net model is tested on the MT dataset, which is frequently seen in academia.

The experimental findings of the generated model have a validation accuracy score of 0.9893, a validation mean IoU score of 0.8128, and a validation MAE score of 0.0099.

the rest of the paper is organized as follows. Section 2 describes the proposed ARdiSE-U-net architecture; section 3 discusses the implementation details and experimental results, and section 4 presents the conclusion.

2. THE PROPOSED METHOD

This paper proposes an attention residue expansion compression and excitation unet (ARdiSE-U-net) architecture for pixel-level DSD.

2.1. Attention Block

The attention mechanism used in this study allows the features from the input and transition tensors to be rescaled using the attention map. The input and transition tensors are reduced to intermediate dimensions by 1×1 convolutions. An attention map is obtained by applying sigmoid activation to these collected features. Important features are highlighted by multiplying this map with the input tensor. Mathematically, the mechanism is expressed as follows:

$$\theta_x = W_\theta * x \quad (1)$$

$$\phi_g = W_\phi * g \quad (2)$$

$$f = ReLu(\theta_x + \phi_g) \quad (3)$$

$$\psi_f = \sigma(W_\psi * f) \quad (4)$$

$$\hat{x} = x.\psi_f \quad (5)$$

where:

θ_x: x transformed into intermediate features

ϕ_g: g transformed into intermediate features

W_θ ve W_ϕ : 1x1 convolution weights

f: Activated intermediate feature

σ: Sigmoid function

W_ψ: Weights forming the attention map

ψ_f: [0,1] Attention map with values in the range [0,1]

Figure 1 shows the architecture of the attention block used.

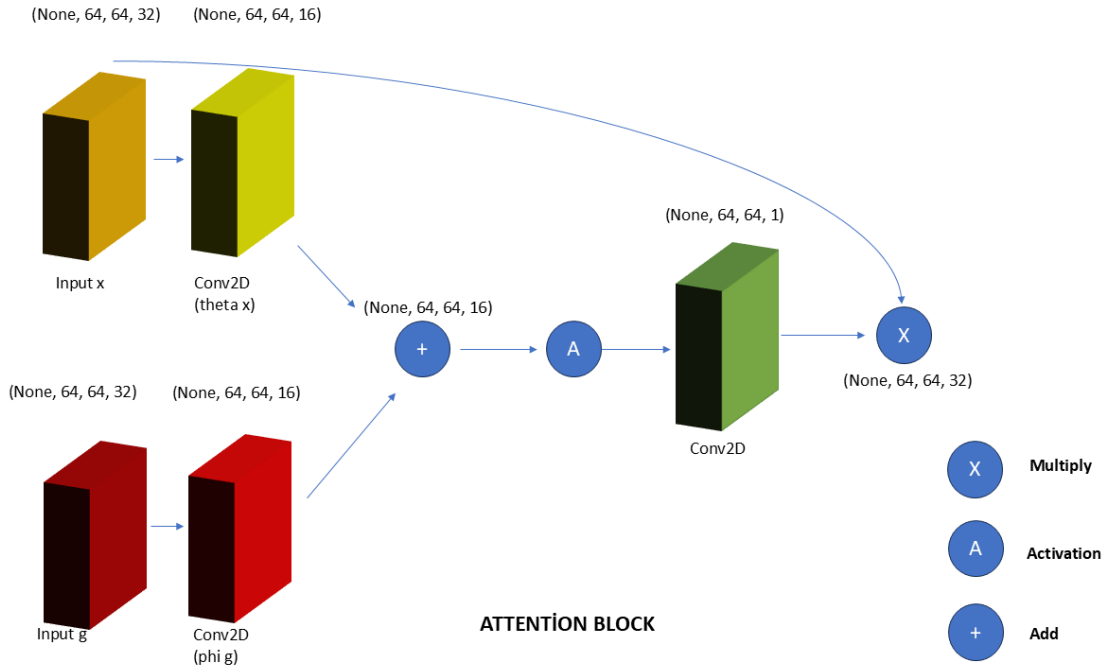


Figure 1. Attention block architecture

2.2. Residual Block with Dilated Convolution

This residual block enables more efficient feature learning using dilated convolution and skip connection. Skip connection allows gradients to propagate more efficiently by preserving the information flow between input and output. Furthermore, the use of dilated convolution expands the detection area without increasing the kernel size. It can be expressed by the following mathematical equations.

$$s = BN(W_s * x) \quad (6)$$

$$y_1 = ReLu(BN(W_1 * x)) \quad (7)$$

$$y_2 = BN(W_2 * y_1) \quad (8)$$

$$y = y_2 + s \quad (9)$$

$$\hat{y} = ReLu(y) \quad (10)$$

where:

x: Input tensor

W_s: 1x1 convolution weights

BN: Batchnormalization

s: Converted input (Shortcut)

W_1: Weights of the first 3x3 extended convolution

W_2: Weights of the second 3x3 extended convolution

$ReLu$: Activation function

Figure 2 shows the architecture of Residual Block with Dilated Convolution.

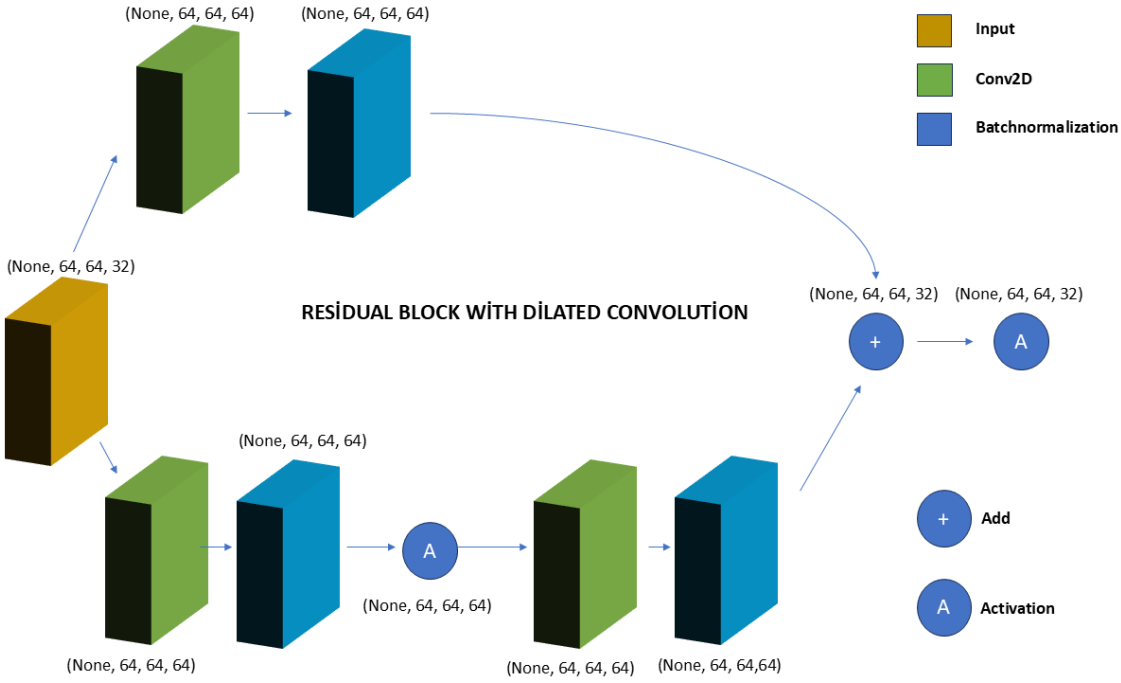


Figure 2. Residual block with dilated convolution architecture

2.3. Squeeze and Excitation Block

Squeeze-and-Excitation (SE) Block is a mechanism that enables the rescaling of channel sizes within a deep learning model using the attention mechanism. It is introduced in the paper "Squeeze-and-Excitation Networks" [31]. SE Block consists of three basic steps: Squeeze, Excitation, and Rescaling.

2.3.1. Squeeze Layer

In this step, the statistical information of each channel in the spatial dimension is compressed, taking into account the channel dimensions. This is performed using Global Average Pooling (GAP). For each channel c, the compressed value is calculated as follows:

$$z_c = \frac{1}{H.W}\sum_{i=1}^{H}\sum_{j=1}^{W} X_c(i,j) \qquad (11)$$

where:

$X_c(i,j)$: pixel value at position *(i,j)* in channel c.

H: Height of the input

W: Width of the input

z_c: compression representation of channel c

There will be a global feature vector $z = [z_1, z_2, \ldots., z_c]$ for all channels of the inputs.

2.3.2. Excitation Layer

In this stage, the compression result z is passed through an attention network to rescale the channel dimensions. Two fully connected layers and an activation function are used:

$$s = \sigma(W_2\delta(W_1 z)) \quad (12)$$

where:

$W_1 \epsilon R^{\frac{C}{r}xC}$: Weights of the first fully correlated layer (compression)

$W_2 \epsilon R^{Cx\frac{C}{r}}$: Weights of the second fully correlated layer (Expansion)

r: Compression ratio

δ: ReLu activation function

σ: Sigmoid activation function

$s \epsilon \mathbb{R}^C$: channel weight vector

This process produces a vector s for rescaling the channel dimensions using the attention network.

2.3.3. Scale Layer

At this stage, the resulting s vector is multiplied by the input tensor (X). This multiplication increases or decreases the importance of each channel in the original tensor:

$$\widehat{X_c} = s_c . X_c \quad (13)$$

where:

$\widehat{X_c}$: Rescaled channel

s_c: Weight for channel c

X_c: The c-th channel of the input tensor

SE Block is usually added in CNN-based models. SE Block is applied to the feature maps resulting from filters in a CNN layer. This process allows the network to focus more on certain features by adding channel-based attention. It also increases the importance of certain features by adding a channel-based attention mechanism. SE Block adds a small computational load and has a lightweight structure. Its shape is given in Figure 3.

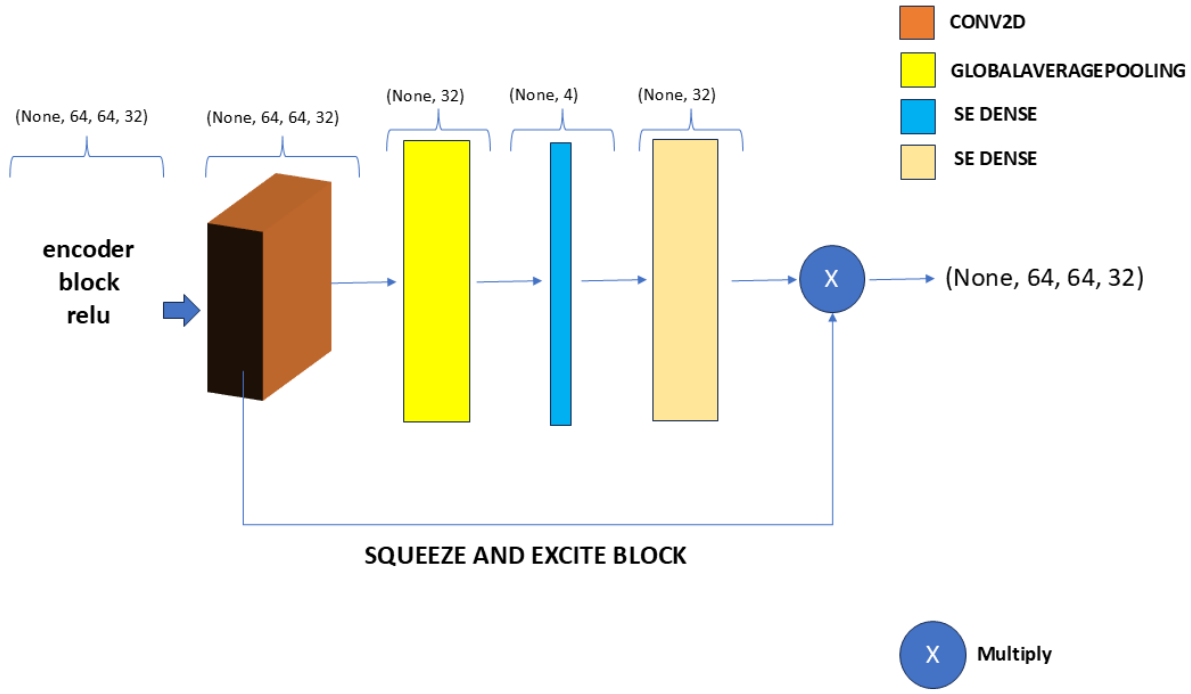


Figure 3. Squeeze and excitation block architecture

2.4. ARdiSE-U-net Architecture

The U-Net model is widely used in deep learning-based image segmentation due to its architecture and effective customization possibilities. In this paper, we describe a model created by applying multiple improvements to the classical U-Net architecture. The model integrates attention mechanisms, residual blocks, dilation applications, and compression-aware (SE) blocks. This combination aims at improving segmentation performance and enabling deeper feature extraction. This architecture consists of several stages. These are the encoder layer, bottleneck layer, decoder layer, and output layer. In the encoder layer, residual blocks and dilated convolutions are used to extract important features from the image. SE blocks highlight important information in the feature maps by applying spatial and channel-based attention mechanisms. In the bottleneck layer, larger filter sizes and widening ratios are used for deep feature extraction. In the decoder layer, attention blocks are added with up-sampling operations. These blocks enable more precise segmentation by utilizing the feature maps at the corresponding encoder level.

In the output layer, pixel-level output is obtained by using a 1x1 convolution layer with sigmoid activation for segmentation. In the output layer, dice loss, the loss function whose equation is given below, is used.

$$Dice\ Loss = 1 - \frac{2.|A \cap B| + \epsilon}{|A| + |B| + \epsilon} \tag{14}$$

This model aims to achieve better results in image segmentation by enriching the powerful structure of U-Net with modern deep learning components. In particular, the integration of attention mechanisms and SE blocks increases the focusing and learning capacity of the model. Expansion and residual blocks improve context information and learning stability. Figure 4, Figure 5 and Figure 6 illustrate the proposed architecture.

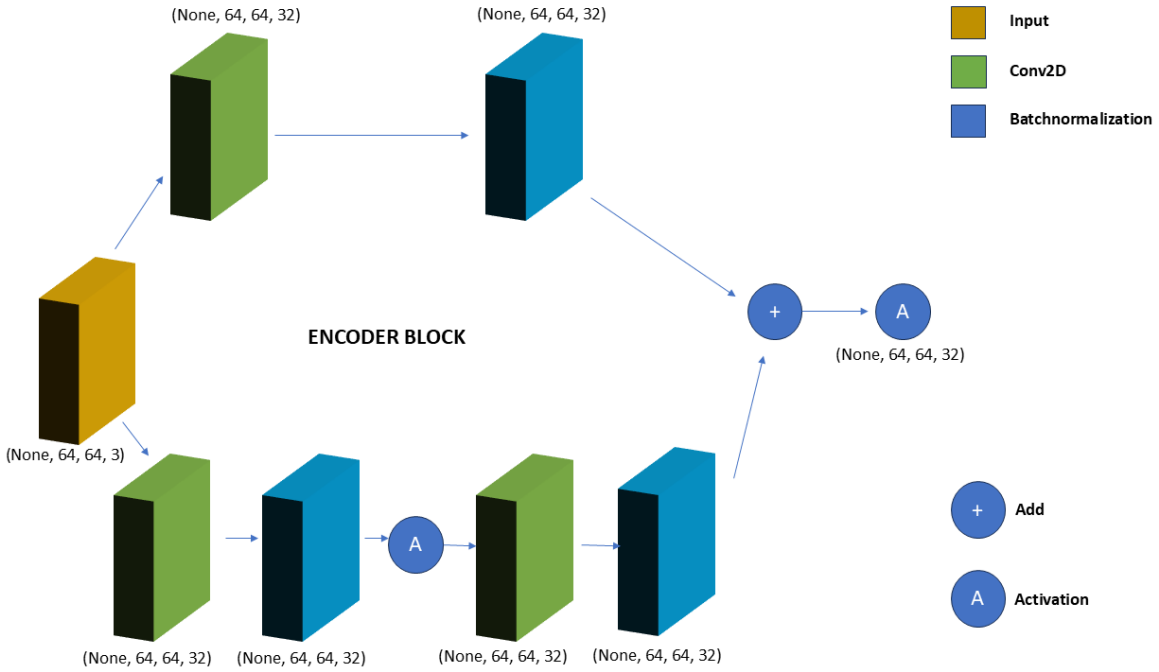


Figure 4. Encoder block architecture

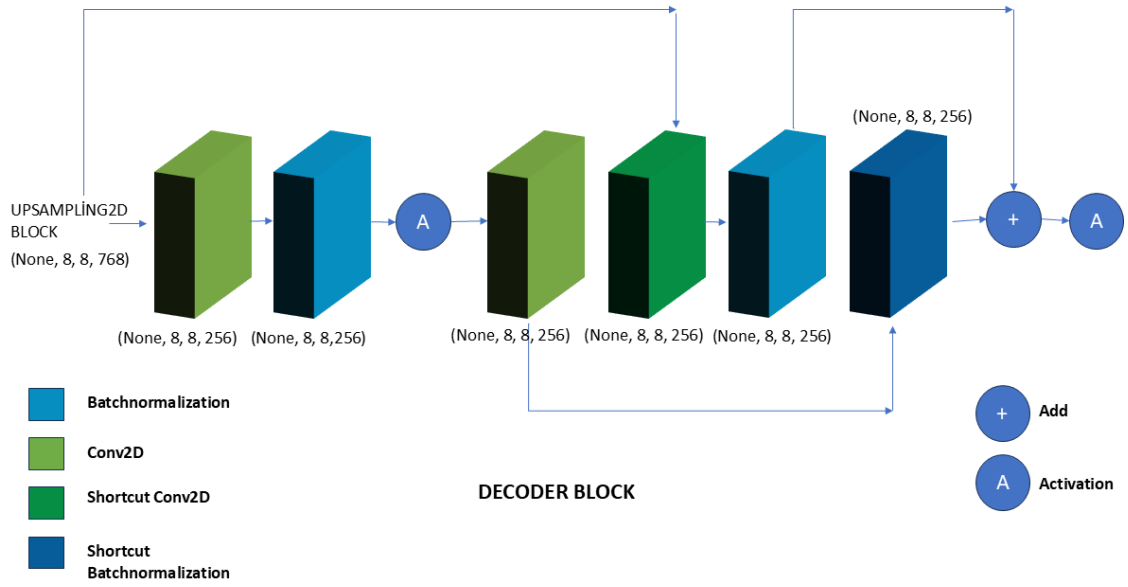


Figure 5. Decoder block architecture

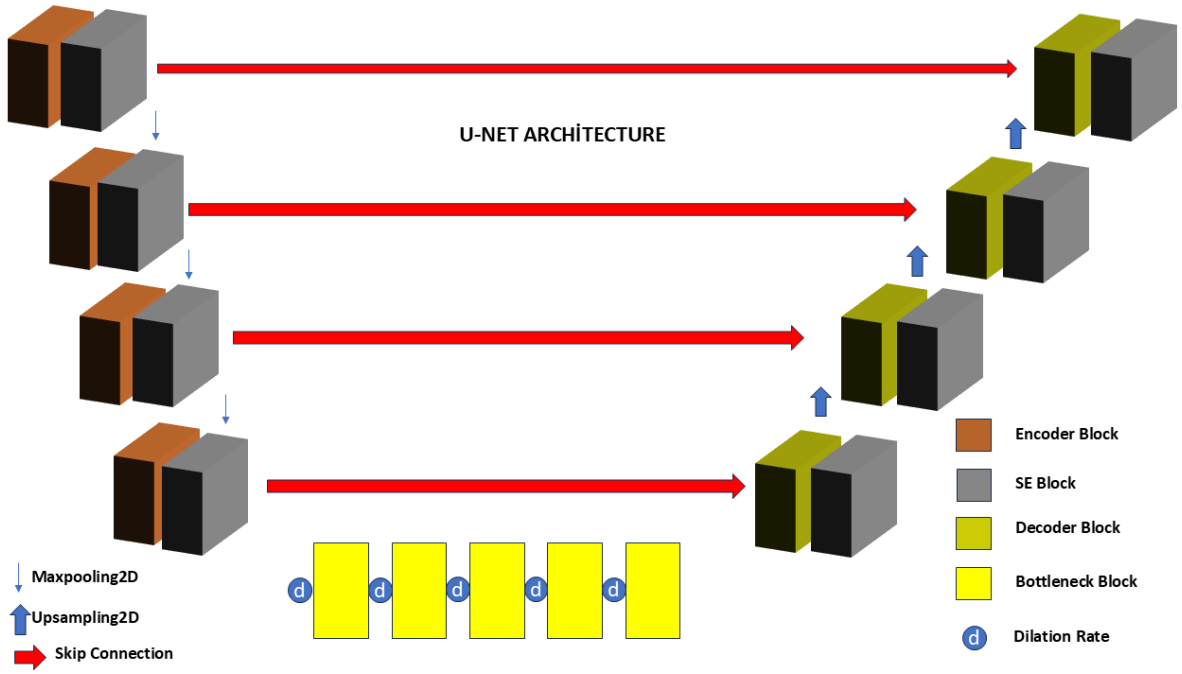


Figure 6. U-net architecture

2.5. Proposed ARdiSE-U-net Training Procedure

In deep learning studies involving methods such as segmentation and classification methods, the effect of the size of the data set on performance is very important. Overfitting or underfitting problems are frequently encountered in small data sets. In order to eliminate these problems, transfer learning or data augmentation (DA) methods are used [24, 33]. DA methods include methods such as flipping, mirroring, cropping, expanding, and playing with color tones. In this study, in order to increase the effectiveness of the proposed ARdiSE-U-net method, only inversion is used.

3. IMPLEMENTATION DETAILS

The dataset we used in the experimental study is [23]. This dataset was used to DSD at the pixel stage and to compare the performance of the based methods with the proposed model. The proposed model in the study was developed and implemented on the Python platform using TensorFlow-Keras libraries. The variables that were utilized during the training step of the system we created are listed below:

- The batch size is 16.

- Learning rate is 0.001.

- Loss function is Dice Loss function

All stages of all models used in this article were run with identical settings. Also, during the experimental research, mIoU (Mean Intercept over Unity), accuracy score, AUC (Area Under ROC Curve), and MAE (Mean Absolute Error) metrics were used to observe the performance of the developed model.

3.1. Dataset

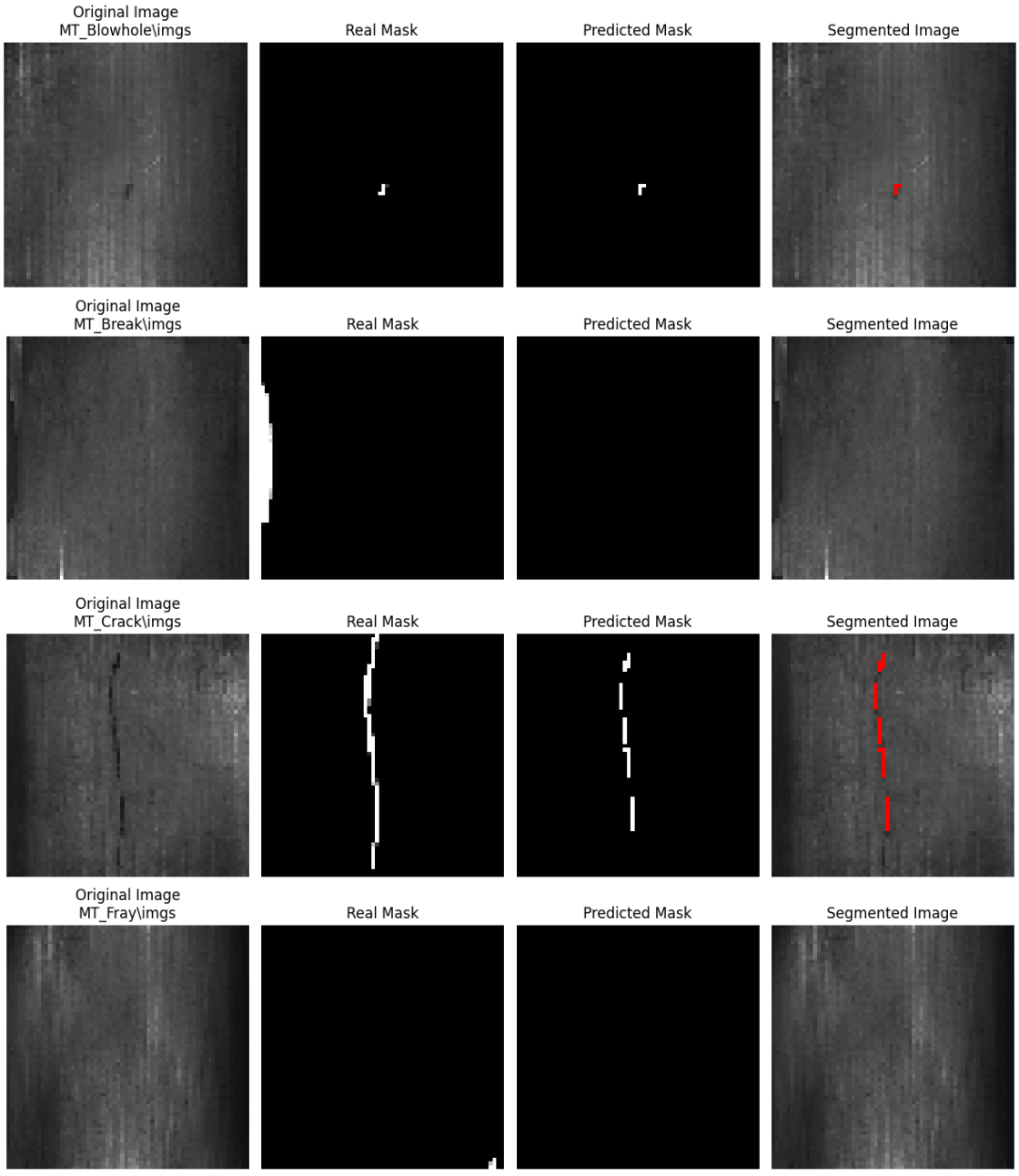
In this study, the MT [23] dataset was used to evaluate the model performances. There are 1344 images in total in 5 different types on magnetic tile surfaces in the MT dataset. All images from the MT dataset are utilized in experiments. In experimental studies, the training and testing rates of the MT dataset are 80% and 20%,

respectively. Additionally, every image in the dataset was scaled to 64 x 64 pixels. DA was applied to the dataset in order to improve model success rates. Also, the data was translated and doubled.

3.1. Experimental Results

The system we used in our study has 16 GB RAM and 6 GB GTX1060 graphics card. The experimental results and graphs in this study are presented below.

The images in the MT dataset have small surface defects and were recorded under different lighting conditions. Therefore, the estimation and DSD are difficult. Figure 7 shows the performance of our proposed architecture on the MT dataset with the original image, real mask, predicted mask, and segmented images.



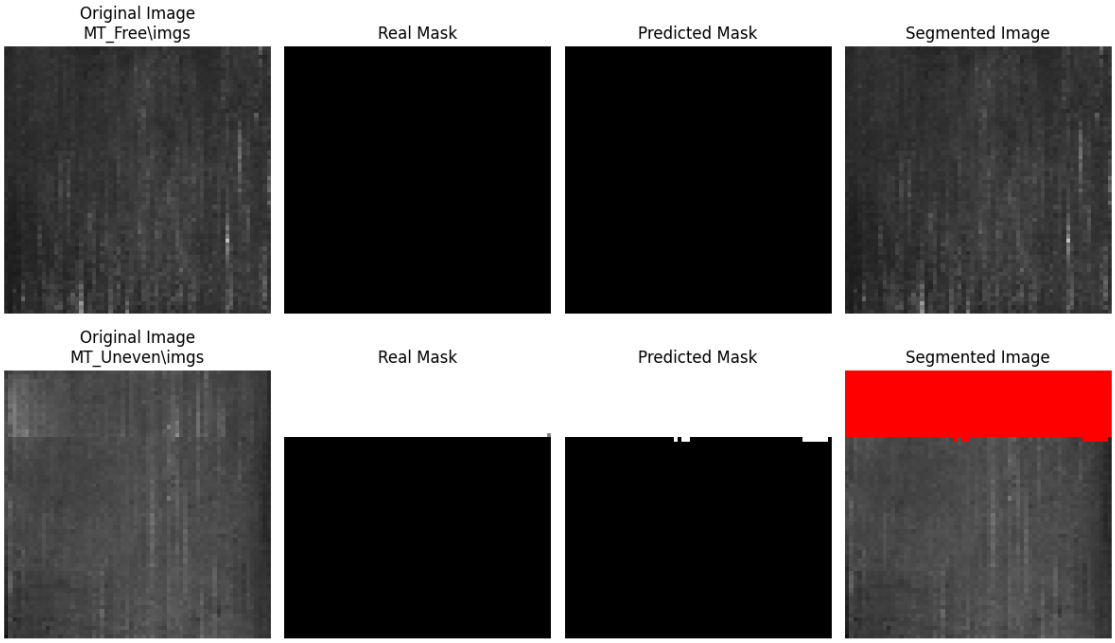
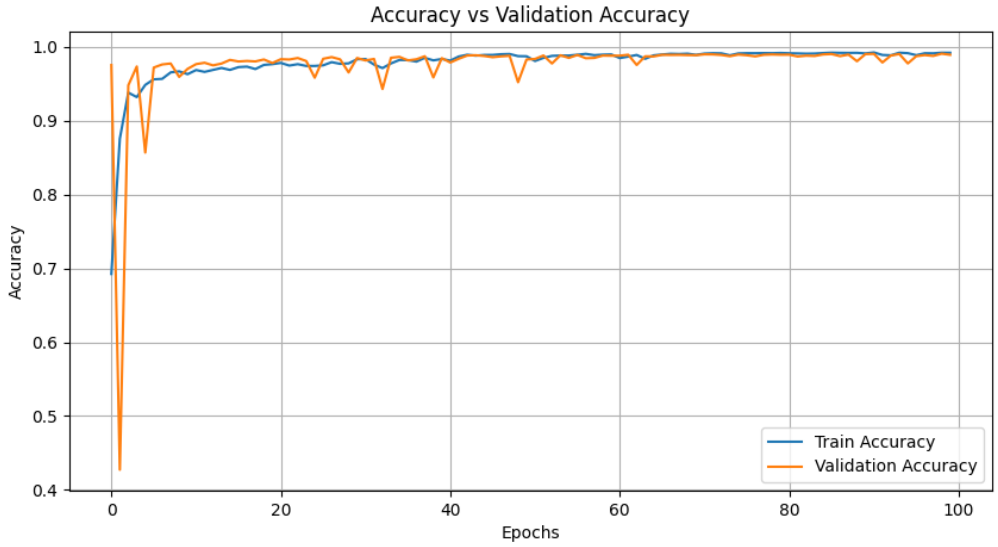


Figure 7. Original image, real mask, predicted mask, and segmented images of our proposed architecture

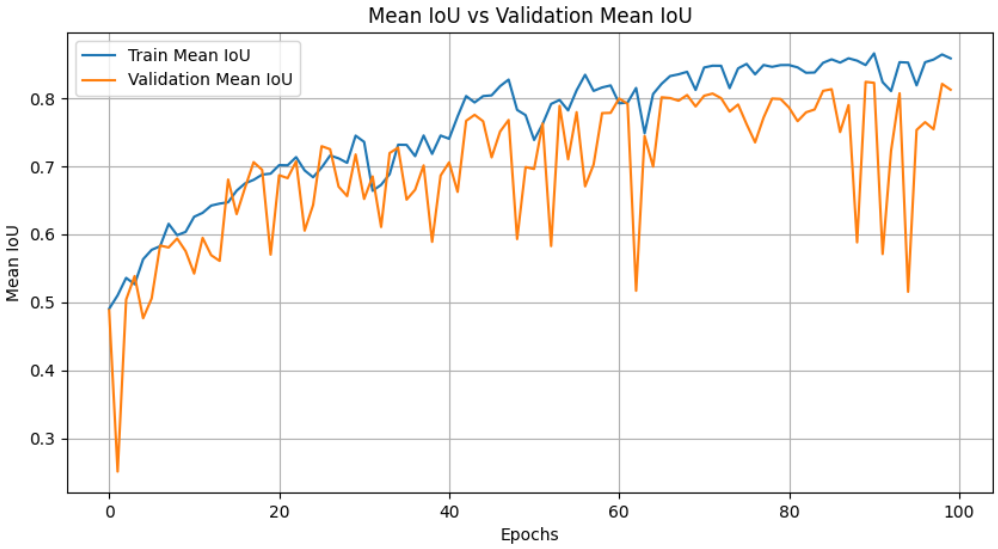
Our proposed architecture preserves the important features obtained from small areas and carries these features to the last layer. In this way, the developed model has high performance compared to most of the baseline models. Furthermore, the Attention-Residual-dilation network built to reinforce the characteristics in Unet's skip connection has exhibited low performance only in lines 2 and 4 of Figure 7. Figure 7 shows successful segmentation results in the remaining rows. Table 1 below shows the accuracy, mean IoU, AUC, and MAE metric results of the test results. In addition, accuracy, mean IoU, ROC, MAE, and loss graphs are also shared in Figure 8(a-e), respectively.

Table-1. Model performance comparison on literature

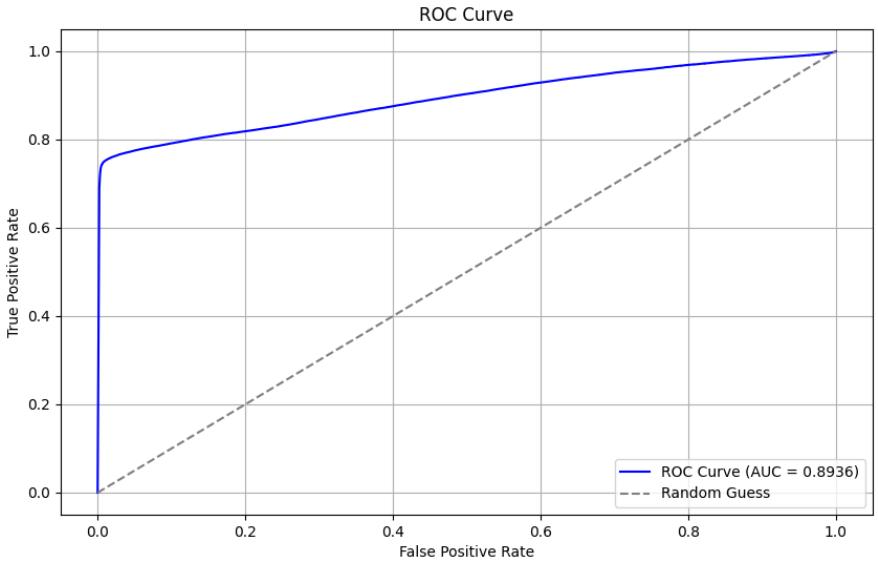
Methods	Accuracy	Mean IoU	AUC	MAE
[34]	-	0.8051	-	-
[23]	-	-	0.985	0.002
[35]	-	0.6694	-	-
[36]	-	-	-	0.0240
[28]	-	0.7131	-	
[27]	-	0.7370	-	--
[37]	-	-	-	-
Proposed method	0.9893	0.8128	0.8936	0.0099



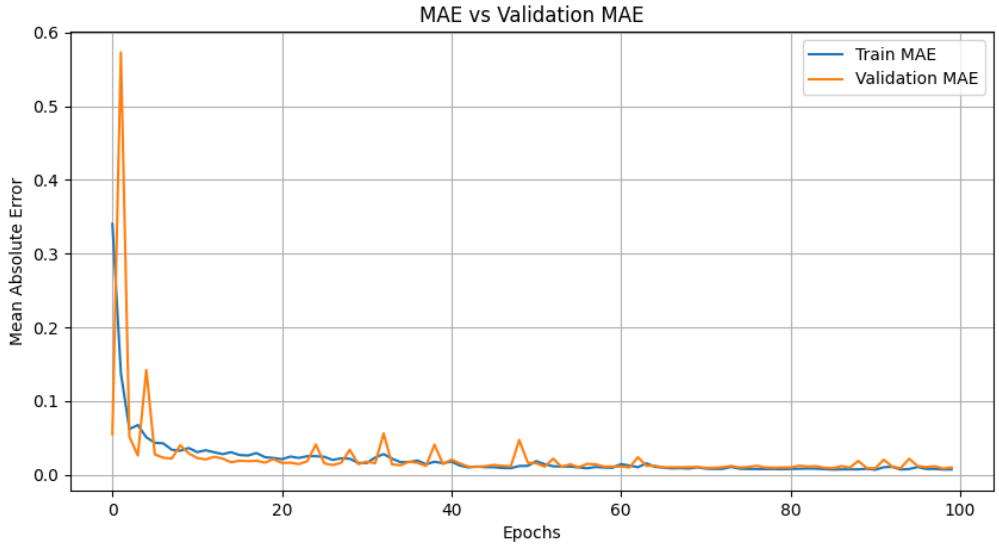
(a)



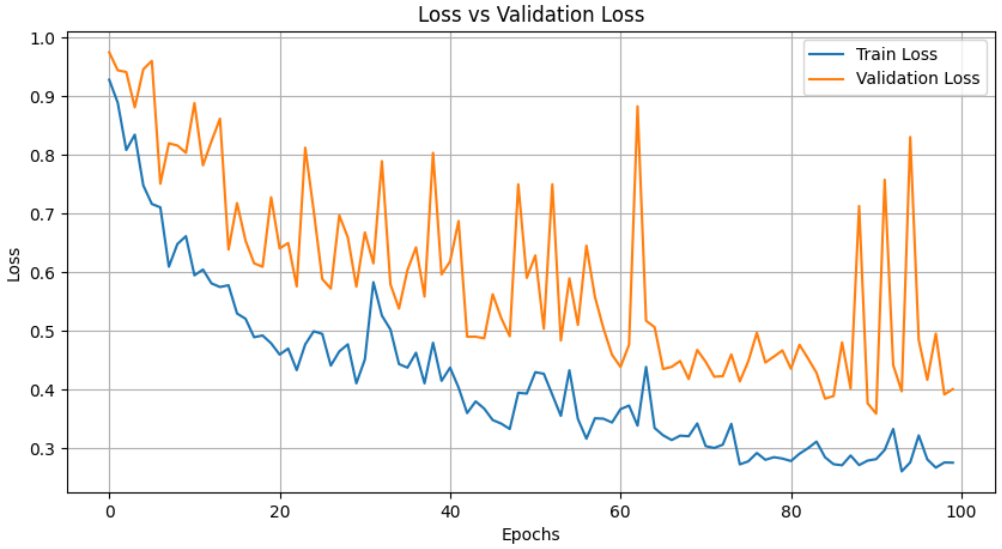
(b)



(c)



(d)



(e)

Figure 8. Accuracy, mean IoU, ROC, MAE, and Loss plots of the developed model

When Table 1 and Figure 8 are analyzed, and the given metrics are considered, the best result in accuracy, mean IoU metrics, belongs to our proposed model. This proves the effectiveness of our proposed model in segmentation and magnetic surface defect detection.

4. CONCLUSION

In this study, ARdiSE-U-net is proposed for pixel-stage magnetic SDS. The basic structure of the developed architecture includes Depth Based Compression and Excitation Block between encoder and decoder skip connections of Unet network. In the model, the detail elements extracted from encoder block are transferred to decoder block while critical details are improved for DSD. Attention-residual layers have also been added. The attention mechanism allows the model to learn that certain regions of an image are more important. Especially in pixel-level tasks such as segmentation, the model needs to process different regions in different ways. It emphasizes important regions of the image (e.g. boundaries of objects or regions of interest). It reduces the influence of unimportant or distracting regions. It is especially helpful in distinguishing between different classes with similar colors and textures. Residual connections have been used in deep learning models to solve the vanishing gradient problem and facilitate the model to learn deeper layers. In segmentation tasks, it allows the information learnt in previous layers to be carried to the next layers. It helps the model to learn faster and

more consistently. It also ensures that information from earlier layers is preserved, especially in segmentation tasks where fine details are important. The Attention Residual layer optimizes both attention weighting and information flow by combining the attention mechanism and residual connections. Experimental results show that our proposed architecture generally produces better results than the baseline studies. This work is expected to shed light on future studies. In this study, a new model, ARdiSE-U-Net, is introduced for the pixel-level DSD. In the proposed architecture, the performance is improved by adding Depth-Based Compression and Excitation Block to the connections between the encoder and decoder layers of the U-Net network. Experiments have shown that this architecture generally provides better results than previous methods. The aim of the study is to provide a guiding basis for future research.

REFERENCES

[1] Yang C, Liu P, Yin G, Jiang H, Li X. Defect detection in magnetic tile images based on stationary wavelet transform. Ndt & E International. 2016;83:78-87. doi:10.1016/j.ndteint.2016.04.006

[2] Li X, Jiang H, Yin G. Detection of surface crack defects on ferrite magnetic tile. Ndt & E International. 2014;62:6-13. doi:10.1016/j.ndteint.2013.10.006

[3] Xie L, Lin L, Yin M, Meng L, Yin G. A novel surface defect inspection algorithm for magnetic tile. Applied surface science. 2016;375:118-26. doi:10.1016/j.apsusc.2016.03.013

[4] Yang C, Liu P, Yin G, Wang L. Crack detection in magnetic tile images using nonsubsampled shearlet transform and envelope gray level gradient. Optics & Laser Technology. 2017;90:7-17. doi:10.1016/j.ndteint.2016.04.006

[5] Üzen H, Firat H, Hanbay D, editors. Automatic thresholding method developed with entropy for fabric defect detection. 2019 International Artificial Intelligence and Data Processing Symposium (IDAP); 2019: IEEE. doi:10.1109/idap.2019.8875890

[6] Sari-Sarraf H, Goddard JS, editors. Vision system for on-loom fabric inspection. 1998 IEEE Annual Textile, Fiber and Film Industry Technical Conference (Cat No 98CH36246); 1998: IEEE. doi:10.1109/texcon.1998.679228

[7] Kaddah W, Elbouz M, Ouerhani Y, Alfalou A, Desthieux M. Automatic darkest filament detection (ADFD): a new algorithm for crack extraction on two-dimensional pavement images. The Visual Computer. 2020;36(7):1369-84. doi: 10.1007/s00371-019-01742-2

[8] Mak K, Peng P, Lau H, editors. Optimal morphological filter design for fabric defect detection. 2005 IEEE international conference on industrial technology; 2005: IEEE. doi:10.1109/icit.2005.1600745

[9] Mingde B, Zhigang S, Yesong L. Textural fabric defect detection using adaptive quantized gray-level co-occurrence matrix and support vector description data. Information Technology Journal. 2012;11(6):673. doi:10.3923/itj.2012.673.685

[10] Hanbay K, Talu MF, Özgüven ÖF. Fabric defect detection systems and methods—A systematic literature review. Optik. 2016;127(24):11960-73. doi:10.1016/j.ijleo.2016.09.110

[11] Bissi L, Baruffa G, Placidi P, Ricci E, Scorzoni A, Valigi P. Automated defect detection in uniform and structured fabrics using Gabor filters and PCA. Journal of Visual Communication and Image Representation. 2013;24(7):838-45. doi:10.1016/j.jvcir.2013.05.011

[12] Cha YJ, Choi W, Büyüköztürk O. Deep learning-based crack damage detection using convolutional neural networks. Computer-Aided Civil and Infrastructure Engineering. 2017;32(5):361-78. doi:10.1111/mice.12263

[13] Masci J, Meier U, Ciresan D, Schmidhuber J, Fricout G, editors. Steel defect classification with max-pooling convolutional neural networks. The 2012 international joint conference on neural networks (IJCNN); 2012: IEEE. doi:10.1109/ijcnn.2012.6252468

[14] Soukup D, Huber-Mörk R, editors. Convolutional neural networks for steel surface defect detection from photometric stereo images. International symposium on visual computing; 2014: Springer. doi:10.1007/978-3-319-14249-4_64

[15] Dai W, Erdt M, Sourin A. Detection and segmentation of image anomalies based on unsupervised defect reparation. The Visual Computer. 2021;37(12):3093-102. doi:10.1007/s00371-021-02257-5

[16] He Y, Song K, Meng Q, Yan Y. An end-to-end steel surface defect detection approach via fusing multiple hierarchical features. IEEE transactions on instrumentation and measurement. 2019;69(4):1493-504. doi:10.1109/tim.2019.2915404

[17] Pastor-López I, Sanz B, Tellaeche A, Psaila G, de la Puerta JG, Bringas PG. Quality assessment methodology based on machine learning with small datasets: Industrial castings defects. Neurocomputing. 2021;456:622-8. doi:10.1016/j.neucom.2020.08.094

[18] Ronneberger O, Fischer P, Brox T, editors. U-net: Convolutional networks for biomedical image segmentation. Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18; 2015: Springer. doi:10.1007/978-3-319-24574-4_28

[19] Lin T-Y, Dollár P, Girshick R, He K, Hariharan B, Belongie S, editors. Feature pyramid networks for object detection. Proceedings of the IEEE conference on computer vision and pattern recognition; 2017. doi:10.1109/cvpr.2017.106

[20] Augustauskas R, Lipnickas A. Improved pixel-level pavement-defect segmentation using a deep autoencoder. Sensors. 2020;20(9):2557. doi:10.3390/s20092557

[21] Long J, Shelhamer E, Darrell T, editors. Fully convolutional networks for semantic segmentation. Proceedings of the IEEE conference on computer vision and pattern recognition; 2015. doi:10.1109/cvpr.2015.7298965

[22] Damacharla P, Rao A, Ringenberg J, Javaid AY, editors. TLU-net: a deep learning approach for automatic steel surface defect detection. 2021 International Conference on Applied Artificial Intelligence (ICAPAI); 2021: IEEE. doi:10.1109/icapai49758.2021.9462060

[23] Huang Y, Qiu C, Yuan K. Surface defect saliency of magnetic tile. The Visual Computer. 2020;36(1):85-96. doi:10.1007/s00371-018-1588-5

[24] Jing J, Wang Z, Rätsch M, Zhang H. Mobile-Unet: An efficient convolutional neural network for fabric defect detection. Textile Research Journal. 2022;92(1-2):30-42. doi:10.1177/0040517520928604

[25] Luo Q, Gao B, Woo WL, Yang Y. Temporal and spatial deep learning network for infrared thermal defect detection. Ndt & E International. 2019;108:102164. doi:10.1016/j.ndteint.2019.102164

[26] Zhang D, Song K, Xu J, He Y, Niu M, Yan Y. MCnet: Multiple context information segmentation network of no-service rail surface defects. IEEE Transactions on Instrumentation and Measurement. 2020;70:1-9. doi:10.1109/tim.2020.3040890

[27] Cao J, Yang G, Yang X. A pixel-level segmentation convolutional neural network based on deep feature fusion for surface defect detection. IEEE Transactions on Instrumentation and Measurement. 2020;70:1-12. doi:10.1109/tim.2020.3033726

[28] Dong H, Song K, He Y, Xu J, Yan Y, Meng Q. PGA-Net: Pyramid feature fusion and global context attention network for automated surface defect detection. IEEE Transactions on Industrial Informatics. 2019;16(12):7448-58. doi:10.1109/tii.2019.2958826

[29] Fu X, Li K, Liu J, Li K, Zeng Z, Chen C. A two-stage attention aware method for train bearing shed oil inspection based on convolutional neural networks. Neurocomputing. 2020;380:212-24. doi:10.1016/j.neucom.2019.11.002

[30] Howard AG. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:170404861. 2017. doi:10.20944/preprints202411.2377.v1

[31] Hu J, Shen L, Sun G, editors. Squeeze-and-excitation networks. Proceedings of the IEEE conference on computer vision and pattern recognition; 2018. doi:10.1109/cvpr.2018.00745

[32] Roy AG, Navab N, Wachinger C. Recalibrating fully convolutional networks with spatial and channel "squeeze and excitation" blocks. IEEE transactions on medical imaging. 2018;38(2):540-9. doi:10.1109/tmi.2018.2867261

[33] Natarajan V, Hung T-Y, Vaikundam S, Chia L-T, editors. Convolutional networks for voting-based anomaly classification in metal surface inspection. 2017 IEEE International Conference on Industrial Technology (ICIT); 2017: IEEE. doi:10.1109/icit.2017.7915495

[34] Üzen H, Turkoglu M, Aslan M, Hanbay D. Depth-wise Squeeze and Excitation Block-based Efficient-Unet model for surface defect detection. The Visual Computer. 2023;39(5):1745-64. doi:10.1007/s00371-022-02442-0

[35] Huang Y, Huang Z, Jin T. DEU-Net: A Multi-Scale Fusion Staged Network for Magnetic Tile Defect Detection. Applied Sciences. 2024;14(11):4724. doi:10.3390/app14114724

[36] Jiang C, Zhang X, Xu B, Zheng Q, Li Z, Zhang L, et al. MT-U2Net: Lightweight detection network for high-precision magnetic tile surface defect localization. Materials Today Communications. 2024;41:110480. doi:10.1016/j.mtcomm.2024.110480

[37] Liu J, Song K, Feng M, Yan Y, Tu Z, Zhu L. Semi-supervised anomaly detection with dual prototypes autoencoder for industrial surface inspection. Optics and Lasers in Engineering. 2021;136:106324. doi:10.1016/j.optlaseng.2020.106324

