

Deep Learning Based Semantic Segmentation Technique for Anomaly Detection on Metal Surfaces Using High Calibre U- Shaped Network

Reshma P Vengaloor, Roopa Muralidhar*

Department of Electronics and Communication Engineering, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Ramapuram, Chennai 600089, India

Corresponding Author Email: roopam@srmist.edu.in

<https://doi.org/10.18280/ts.390614>

ABSTRACT

Received: 24 August 2022

Accepted: 15 December 2022

Keywords:

anomalies, deep learning, high calibre U-net, semantic segmentation

Automatic detection of anomalies on the metal surface is an essential capability in industries to provide the better-quality control. To locate and identify the type of defect, it is necessary to find the Region of interest (RoI) from the captured image. Segmentation of the captured image is one among the many methods to achieve this task. Therefore, a precise and accurate segmentation method has major role to improve the metal surface anomaly detection rate in industry. As the defects are different in its size, shape and type, the process of semantic segmentation for metal surface is considered as a challenging task. To address this issue, a deep learning based high calibre U- shaped network is proposed. It can be considered as an automatic quality control system for industries. The proposed method is effective in predicting the presence of defects. The system is also capable to locate the position of the defect on surface without the intervention of human being. The up-sampling technique provided with the convolutional neural network in the architecture makes the system to produce high resolution outputs. The proposed system has been evaluated based on accuracy, precision, loss and IoU after training and testing the model using two different datasets called NEU metal surface defect database and Kolektor surface defect data set.

1. INTRODUCTION

Ensuring the better-quality control in industry where the products are getting shipped is very much important. The inspection of the surface of a finished product is the way to assure the quality. The appearance and the quality will get affected if there any damage on the surface. Depending on the extend or the type of the anomaly or damage, corrective action can be taken to reject the product or maintain it. Quite often, such kinds of surface quality control processes are carried out manually with the help of workers. However, this process will consume lot of time and at the same time it leads to the inefficient production or missed defect detection.

The image processing techniques called edge detection, thresholding in gray scale images etc. are the few primary steps of surface anomaly detection. The recent technology developments and advances in the area of computer vision can make a remedy for such problems. The development of automatic defect detection systems that can recognize, locate, and categorize faults, which is made possible by the advances in deep learning and image processing technology. Deep learning technologies make accurate predictions, which may be used to increase production quality and detection accuracy in the industrial sector. They will be processing and analyzing the images of the surfaces to check for the defects. There are several methods developed [1-5] typically for the defect detection. For example, blob detection algorithm is explained in paper [6] for detecting the defect on the surface of a tile. There is a possibility of occurrence of several forms of uncertainty in the industrial or manufacturing sectors. The anomalies may be variant in its size, shape or depth.

Consequently, a technique that can adapt several variants must be created. Learning-based methods can thus replace the preprogrammed feature extraction techniques since they are able to provide better robustness to the variations in uncertainties. Learning-based techniques include support vector machines and neural networks. However, they are not designed to be able to recognize patterns in flaws. They are less robust to the variations in the texture and complexity of the defect.

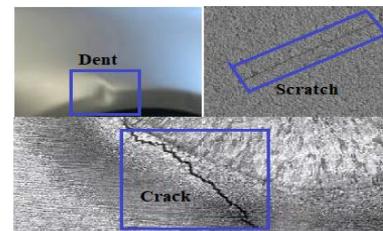


Figure 1. Captured image to show the occurrence of different types of defects on metal surfaces

Deep learning methods have been proved to be robust to background lighting, color, size and shape, which is an advantageous attribute for complex surface inspection in industrial settings. The major challenges in defect detection are the occurrence of defect in various shape, existence of ambiguous edges, low contrast and the occurrence of different numerous backdrop colors and scratches. Figure 1 shows the captured image to indicate various types of defects which can occur over the metal surfaces. Identification of defects like dents, cracks and scratches is concentrated in this paper. The

cause of occurrence of these anomalies are different. Dent will get formed when the surface is being pushed in or hit. Based on the depth of the dents they are categorized as round dents, sharp dents, creased dents, extreme dents and miscellaneous dents. The scratch on the surface is said to be a mark on the surface due to sharp or jagged objects, whereas the cracks are the split on the surface without breaking them apart.

The deep learning-based defect detection is categorized as supervised, semi-supervised, and unsupervised learning. The training dataset in supervised learning will be labelled, wherein unsupervised learning, the labelled data will not be available. Unsupervised learning will be based on the following assumptions: 1) The majority samples available in the dataset are non-defective. 2) The defective instance should show significant variation from the normal instances in the data. It is demonstrated a defect detection based on the unsupervised learning and deep generative adversarial networks (GAW) [7].

An automatic defect detection system should be capable of identifying and classifying the defects. The existence of normal regions along with the defective regions may give poor accuracy [7] in classification. With the evolution of Fully Connected Neural Network (FCNN) in 2015, the state-of-the-art performance on the images become better [8]. Utilizing the advantages of CNN such as high efficiency and accuracy, there have been invented several methods for surface defect detection.

To provide the higher efficiency and accuracy of the metal surface defect detection, semantic segmentation using high calibre UNET is proposed in this paper. The system will find the region of interest by segmenting the pixel variations over the image of metal surface. The proposed system is capable of providing high resolution outputs as it implements the up-sampling operation in the architecture. The up sampling is accomplished by concatenating the features from the encoder path of the architecture.

The whole layout of this paper is given as follows: Section 2 describes the related works and following which the proposed system is explained in section 3. Section 4 is stating the results and discussions. Conclusion and the future directions are expounding in section 5.

2. RELATED WORKS

A non-defect detection without manual intervention is proposed by Wheeler and Karimi [9]. It analyses the images that were taken with the right camera and light source, locates the problem, and then classifies it using the appropriate defect detection method. This method can provide high detection efficiency. The enhancement of surface defect detection technologies was made possible by the breakthrough in the field of machine vision. Ground Penetrating Radars are used in the study of Zhang et al. [10] to analyze the moisture damages on pavement bridges. It has used a mixed deep convolutional neural network for extracting the features. To detect and quantify the defects on the concrete surfaces, the deep learning technology along with the structured light composed of two laser sensors are used in the study of Park et al. [11]. To increase the accuracy in size measurement, the laser alignment correction algorithm is used. The novel defect detection method-based image morphology and Huffman transform is validated in the study of Cao et al. [12]. The Grab Cut algorithm is used [13] with the combination of deep neural

network to provide the skin lesion segmentation. A YOLO-based object detection architecture in combination with image processing algorithm has been explained in the study of Tao et al. [14].

In paper [15], the improved deep learning convolutional neural network is tested for detection of defects, which will reduce the labor cost for manual extraction of defect features. A new classification network using multi-Group convolutional neural network (MG-CNN) is presented in the study of He et al. [16]. It will extract the feature map groups from different types of defects. The method explained in the study of Mandal et al. [17] has collected large number of images as the deep learning techniques need a huge dataset. The accuracy of detection and classification was measured using the F1 score, which was found using the precision and recall values. The labeled training data base is not easily available as it requires domain specific knowledge. Therefore, an active learning method is explained in the study of Dai et al. [18] to reduce the labeling workload. Based on deep learning, a novel supervised method is proposed by Huang et al. [19]. This method will adjust the layers in YOLOv3 architecture with more valuable samples. These samples are selected using a less confident strategy. The method has been proved that the number of labeled images can be reduced to train the model without decreasing the accuracy. A supervised processing pipeline is proposed by Qiao et al. [20]. It requires only image level labeled data rather than the bounding box. Labeled dataset approach has been proposed by Majidifard [21]. In this the data set contains the images, which are captured from two camera views of an identical pavement segment. These wide-view images will be used for training the deep learning model as well as for the classification of the distresses. The dimension clustering of target frames is carried out by Jing et al. [22]. It combines the defect size and K - Means algorithm to reduce the error detection rate. Two-layer detection algorithm is proposed by He et al. [23]. It is capable of handling huge number of images with complex backgrounds. For each layer different feature extractions networks have been used. The results of the tests proved that the two-detection-layer algorithms may perform better than the single-layer detection method.

Comparing with other machine learning methodologies, Convolutional Neural Network (CNN) can provide high accuracy. Therefore, CNN is one of the most used deep learning models for object detection [24]. It is considered as the faster deep learning model without reducing the effectiveness. A new efficient deep learning model is proposed by Jeon and Jeong [25] for text detection. The proposed model combined the technology of MobileNetV2 and a balanced decoder. The balanced decoder is composed of stack of Inverter Residual Block (IRB) and standard convolutional layers. The paper [26] proposed to identify the cracks and potholes on pavements. In this the multi spectral images of the pavements are captured using PAV. Certain machine learning methods, such as SVM, Random Forest, etc., have been used to distinguish between photos of damaged and normal pavement. For monitoring the rigid conditions of pavements, UAV-based crack detection method is proposed by Ersoz et al. [27]. An automated surface crack detection system using Otsu thresholding and morphological operation is proposed by Dorafshan et al. [28]. The paper [29] explains a method for fatigue crack detection using UAV for under bridge inspection system. This also explained an image processing algorithm for the detection of steel fatigue as well as crack fatigue. The

advantage of training and dyeing solutions is not used in the method. Therefore, the system is obtained as quick and effective. A method of crack detection using hybrid image processing has been described in the study of Kim et al. [30]. It comprises a two-step method. The images are captured using UAV and then the histogram analysis will be performed for crack detection. The AR Drone is equipped with a Raspberry Pi, camera, and ultrasonic displacement sensor for collecting the distance data and photos. A method using YOLOv2 model for detecting the cracks on concrete in real time is explained in the study of Murao et al. [31]. In this the UAVs are fitted with video camera for capturing the images from concrete walls. This is a unique method in identifying the cracks on concrete bridges. Another bridge crack detection system is discussed in the study of Parker et al. [32]. It uses a coordinated aerial swarm system. Collision prevention of aerial swarm is achieved using this method. CNN is used to classify the images. An automated crack morphology detection using CNN is described in the study of Kim and Cho [33]. In this the Alex Net architecture is used for identifying the cracks. The real time video for training the model has been captured using drone.

Structural steel damage detection system has been explained in the study of Liu and Zhang [34]. The deep learning model VGG16 (Visual Geometry Group) is used for condition evaluation. The performance of the edge detector and Deep Convolutional Neural Network (DCNN) for concrete damage detection is explained in the study of Dorafshan et al. [35]. The method demonstrated that DCNN is providing better performance for crack detection. It also explains a hybrid approach for minimizing the noise. Semantic segmentation for crack detection is discussed by Huang et al. [36]. It uses a self-developed image capturing system for preparing the training data set and test data set. For structural damage identification, a deep learning method has been introduced in the study of Gao and Mosalam [37]. It uses VGGNet-based transfer learning techniques as well as structural ImageNet for avoiding overfitting. Using a deep fully connected network a concrete crack segmentation method is proposed by Dung [38]. This method classifies 40000 images of public concrete crack datasets using three separate pre-trained network architecture.

3. PROPOSED SYSTEM

The block diagram of the proposed system for semantic segmentation of metal surfaces using high calibre U-shaped network is shown in Figure 2. The input images are taken from Kolektor surface defect dataset and NEU metal surface defect database. The image preprocessing techniques will rescale the images into square shape by removing the extra pixels from left and right sides of the images. Then they will be converted from RGB to gray scale using weighted method.

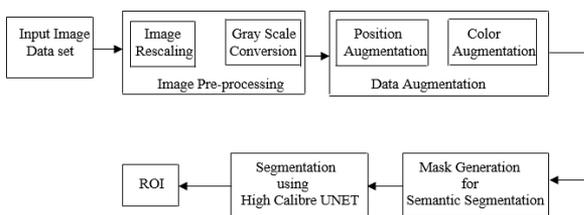


Figure 2. Proposed system for semantic segmentation using high calibre UNET

It is essential in supervised learning to annotate the images in dataset before using them to train the deep learning model. Figure 3 shows an example of a defective metal surface and the corresponding annotation. Image annotation will provide a particular standard, which the deep learning model will follow to make the prediction. Therefore, creating the annotation plays a major role to get appropriate results. The preprocessed images will be passed through data augmentation module as deep learning methods require plenty of images to process. Augmentation is the method which multiplies the number of datasets. This can be achieved through position augmentation and color augmentation methods. Position augmentation can be made by changing the pixel values of an image through the processes like scaling, flipping, padding etc. Changing the color properties like contrast or brightness of an image will provide color augmentation.

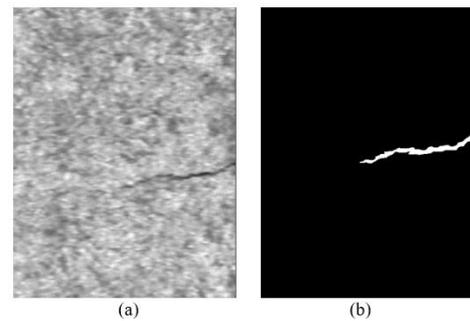


Figure 3. (a) Image of defective metal surface (b) Corresponding annotation

The pixel variation in the image will be considered as the anomaly or the defect on the metal surface. Mask will be generated and used to identify the surface defects. Mask is considered as a filter, which will move through the image to know the variations in pixel value. UNET is the main architecture used in this paper to provide the semantic segmentation. The basic architecture of UNET model is given in Figure 4.

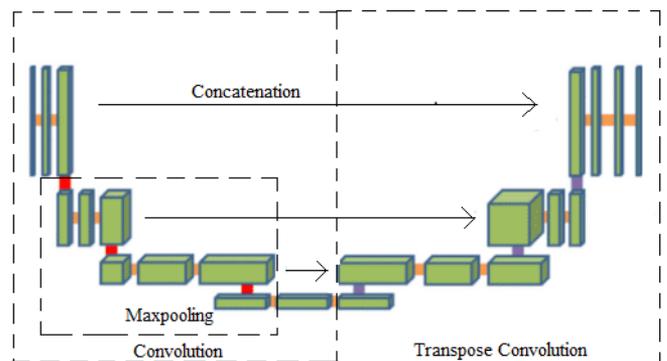


Figure 4. Basic UNET model

UNET is a special type of architecture for image segmentation purpose, where the deep learning tools like convolutional layers and max pool layers will be arranged in such a way to produce the segmented image as output. The typical positioning of convolutional layers and max pool layers are shown in Figure 5.

The first layer is the input layer which has a dimension of $224 \times 224 \times 3$, where 3 indicates that it is a RGB image. The image will be going into the convolutional layers. In the next

layer the dimension has been changed from 3 to 96. Each response will have a dimension of 55 x 55 x 96. Again, a filter of size 5 x 5 is applying. If the filter is moved by one step, then it is said to be stride 1. The number of strides will determine the dimension of the next output layer. The final layers comprise the dense layers.

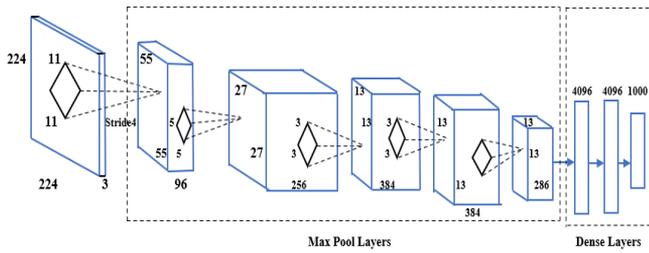


Figure 5. Typical Arrangement of convolutional and max-pool layers

3.1 High calibre U- shaped network (UNET)

3.1.1 Defining and training the architecture model

Figure 6 shows the high calibre UNET architecture for semantic segmentation. It has two paths in its architecture. The path on the left side is known as contraction path or encoder path and the path on the right side is known as expansion or decoder path. The input layer is provided at the top left of the architecture, which has dimension of 128 x 128 x 3. Then the feature size of 16 has been added to get an output of dimension 128 x 128 x16. The kernel size used here is 3 x 3. Padding is provided to treat edge pixels. This process will add extra pixels to the edges, so that the dimension of the output image will be exactly same as the input image. C1 is a two convolutional operation.

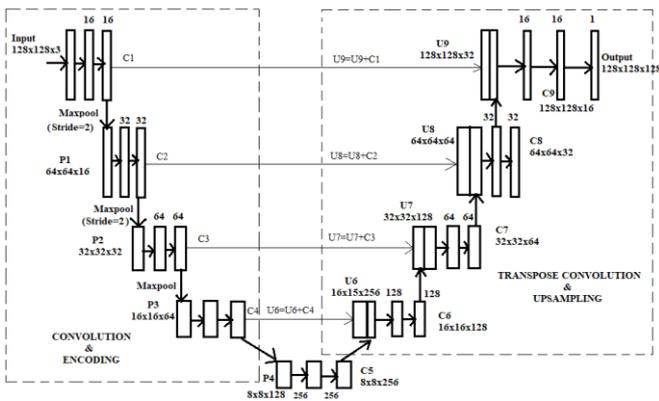


Figure 6. High Calibre UNET architecture for semantic segmentation

P1 is the max pool layer where a 2x2 stride has been used. Within this size the maximum values will be selected and replaced the matrix with the maximum value. As a 2x2 matrix is used, the dimension in P1 will become half of the previously used one. i.e., 64x64x16. Then the feature size of 32 will be given for two convolutions and it produces an output with a dimension of 64 x 64 x 32. The same process will repeat till the last block of the encoder path. Here the dimensions at the input side are 8x8x128 and at the output side is 8x8x256.

For doing up sampling, a process called concatenation is done. After concatenation, the data will pass through two convolutional layer and provide an output of size 16x16x1.

conv2d_transpose_11 (Conv2DTranspose)	(None, 128, 128, 16 2064)	['conv2d_53[0][0]']
concatenate_11 (Concatenate)	(None, 128, 128, 32 0)	['conv2d_transpose_11[0][0]', 'conv2d_39[0][0]']
conv2d_54 (Conv2D)	(None, 128, 128, 16 4624)	['concatenate_11[0][0]']
dropout_26 (Dropout)	(None, 128, 128, 16 0)	['conv2d_54[0][0]']
conv2d_55 (Conv2D)	(None, 128, 128, 16 2320)	['dropout_26[0][0]']
conv2d_56 (Conv2D)	(None, 128, 128, 1) 17	['conv2d_55[0][0]']

 Total params: 1,941,105
 Trainable params: 1,941,105
 Non-trainable params: 0

Figure 7. Number of trainable parameters obtained

Figure 7 shows the total number of parameters obtained after defining the architecture of high calibre UNET. To obtain the total number of trainable parameters, consider the image of size 3x3 and a 2x2 filter. From the example shown in Figure 8, the total number of trainable parameters can be found as 1x(2x2)x3+3=15. Therefore, the number of trainable parameters can be obtained using Eq. (1).

$$\text{Total number of trainable} = \text{Weighted output} + \text{Biase} \quad (1)$$

Parameters

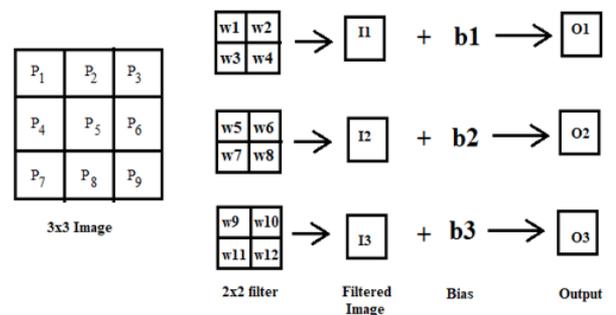


Figure 8. Example for obtaining the trainable parameters

Dropout of 0.1 has been selected to avoid overfitting in the model. Therefore 10% of the neurons will get dropped out. Rectified Linear Unit (ReLU) is used as activation function in defining the high calibre UNET architecture model. At the output layer 'Adam' optimizer has been used. To differentiate the defective portion and non-defective portion, binary cross-entropy is used as it is a binary classification.

3.1.2 Model fitting

The validation split has been chosen as 0.1 to do the model fitting, which means that 10% of the dataset will be used for validation purpose. The performance of the system will be determined based on how the model predict the result when it tests the data from validation dataset. The batch size has been chosen as 16.

4. RESULTS AND DISCUSSIONS

The developed model of high calibre UNET architecture has been tested on Kolektor dataset and NEU data set for metal surface images. Both datasets contain around 650 images

including both damaged as well as normal ones. 70% of the dataset has been chosen for training the proposed model and the remaining 30% has been used for testing the model. The simulation was done using Python programming language by using TensorFlow as basic library module. After defining the CNN layers, all network parameter values have been left as their parameter values. The architecture has been used Adam optimizer.

4.1 Performance metric

The effectiveness of the semantic segmentation provided by high calibre UNET architecture has been evaluated using different parameters such as accuracy, precision and Intersection Over Union (IoU). Accuracy, precision and recall can be defined using Eq. (2), Eq. (3) and Eq. (4) respectively.

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (2)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (4)$$

where, TP refers to True Positive, TN refers to True Negative, FP refers to False Positive and FN refers to False Negative.

Intersection over Union is to find out the similarity between two datasets. In this paper the prediction values and true values are considered as two different datasets to compare. The area of true positives, false negatives and false positives is divided by the area of union will provide the Intersection over Union as shown in Figure 9. In this case true positives are nothing but the number of defects predicted correctly.

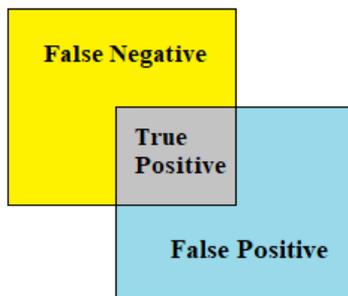


Figure 9. Intersection over union

Considering the ground truth and the prediction of the developed model, the Intersection over Union can be found using Eq. (5).

$$\text{IoU} = \frac{\text{area} (R_g \cap R_p)}{\text{area} (R_g \cup R_p)} \quad (5)$$

where, R_g is the ground truth pixel difference and R_p is the predicted segmentation area.

4.2 Experimental results

The semantic segmentation using high calibre UNET architecture for finding anomalies like dents, cracks and scratches on the metal surfaces has been tested using 30% images from data sets. Figure 10 shows the output obtained when the model run for the image from NEU surface defect

data base and the Figure 11 shows the output for the image from Kolektor data set.

The performance of the developed model is also analysed using various parameters. As it is a semantic segmentation method, the defected and normal area are differentiated using binary classification. i.e., the defective area will be indicated using white colour and the normal area will be indicated using black colour.

The number of epochs has been set to be 125 and the batch size has been given as 16. It is observed that the proposed model has been provided the maximum validation accuracy value of 0.9696 for NEU surface defect dataset as shown in Figure 12. Figure 13 is indicating the graph of training and validation accuracy in each epoch and the Figure 14 is indicating the information on maximum value of accuracy obtained.

The epoch loss is obtained after the execution of each epoch is shown in Figure 15 and epoch loss obtained specifically at epoch 20 is shown Figure 16. It is observed that the value of epoch loss is 0.08130 when the accuracy obtained as maximum. Figure 17 is showing graph for training and validation accuracy when the image from Kolektor data set is running. Epoch loss in training and validation data set is found for the same data set, which is shown in Figure 18.

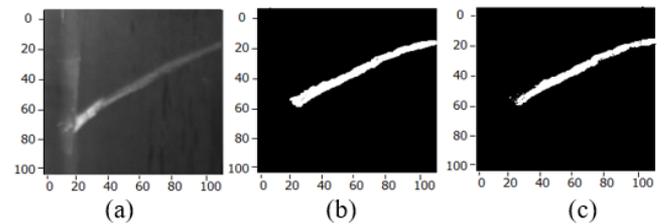


Figure 10. (a) Scratched Image from NEU surface defect database (b) Ground Truth (c) Model Predicted Output

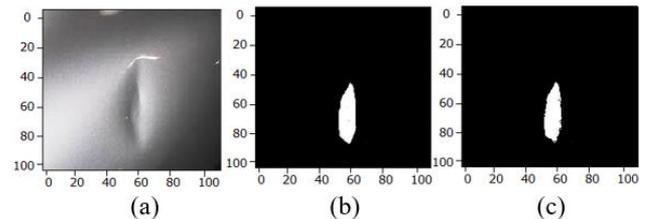


Figure 11. (a) Kolektor surface defect data set image with dent (b) Ground Truth (c) Model Predicted Output

```
Python console
Epoch 15/100
592/603 [=====] - ETA: 0s - loss: 0.0846 - accuracy: 0.9676
Epoch 0015: val_loss did not improve from 0.07900
603/603 [=====] - 24s 39ms/sample - loss: 0.0844 - accuracy: 0.9676 -
val_loss: 0.0794 - val_accuracy: 0.9685
Epoch 16/100
592/603 [=====] - ETA: 0s - loss: 0.0838 - accuracy: 0.9678
Epoch 0016: val_loss improved from 0.07900 to 0.07693, saving model to model_for_surface.h5
603/603 [=====] - 25s 41ms/sample - loss: 0.0832 - accuracy: 0.9680 -
val_loss: 0.0769 - val_accuracy: 0.9689
Epoch 17/100
592/603 [=====] - ETA: 0s - loss: 0.0833 - accuracy: 0.9676
Epoch 0017: val_loss improved from 0.07693 to 0.07540, saving model to model_for_surface.h5
603/603 [=====] - 25s 41ms/sample - loss: 0.0827 - accuracy: 0.9678 -
val_loss: 0.0754 - val_accuracy: 0.9695
Epoch 18/100
592/603 [=====] - ETA: 0s - loss: 0.0799 - accuracy: 0.9690
Epoch 0018: val_loss did not improve from 0.07540
603/603 [=====] - 24s 39ms/sample - loss: 0.0798 - accuracy: 0.9691 -
val_loss: 0.0787 - val_accuracy: 0.9693
Epoch 19/100
592/603 [=====] - ETA: 0s - loss: 0.0775 - accuracy: 0.9699
Epoch 0019: val_loss did not improve from 0.07540
603/603 [=====] - 24s 39ms/sample - loss: 0.0779 - accuracy: 0.9697 -
val_loss: 0.0774 - val_accuracy: 0.9696
Epoch 0020: val_loss did not improve from 0.07540
603/603 [=====] - 24s 39ms/sample - loss: 0.0819 - accuracy: 0.9684 -
val_loss: 0.0862 - val_accuracy: 0.9675
```

Figure 12. Epochs and corresponding validation accuracy obtained for NEU surface defect data set



Figure 13. Training and Validation accuracy obtained for NEU surface defect data set

Name	Smoothed Value	Value	Step	Time	Relative
train	0.9686	0.9684	20	Tue Jul 26, 16:28:30	33m 50s
validation	0.9683	0.9674	20	Tue Jul 26, 16:28:30	33m 50s

Figure 14. Maximum accuracy value obtained during training and validation

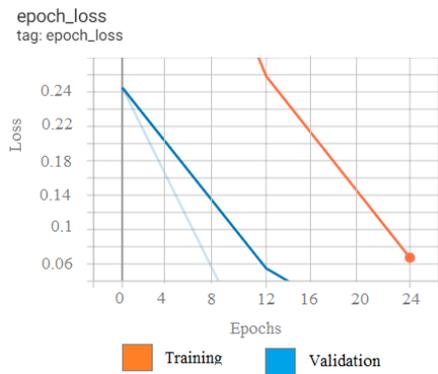


Figure 15. Epoch loss for training and validation data for NEU surface defect database

Name	Smoothed Value	Value	Step	Time	Relative
train	0.08110	0.08190	20	Tue Jul 26, 16:28:30	33m 50s
validation	0.08125	0.08130	20	Tue Jul 26, 16:28:30	33m 50s

Figure 16. The value of epoch loss when the value of accuracy is obtained as maximum



Figure 17. Training and validation accuracy for Kolektor surface defect data set

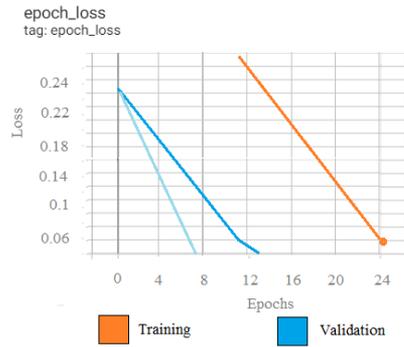


Figure 18. Epoch loss obtained for Kolektor dataset

The performance evaluation of the proposed system, which has been done using NEU metal surface dataset and Kolektor dataset is tabulated in table 1.

Table 1. Performance evaluation on different datasets

Dataset	Accuracy	Loss	IoU
NEU	0.9696	0.08130	0.9132
Kolektor	0.9692	0.08449	0.9124

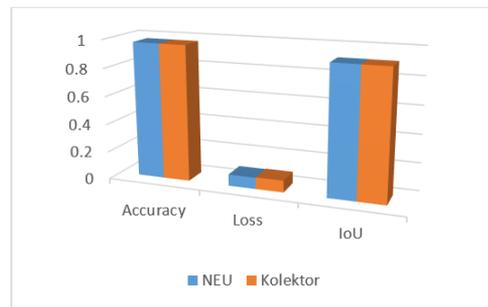


Figure 19. Performance analysis of High Calibre UNET Architecture

The graphical representation performance analysis of high calibre UNET architecture towards various parameters are shown in Figure 19.

Table 2. Comparative analysis

Method	Accuracy	Precision	IoU
Re-weighted Hog & Segmentation	82%	74%	85%
DeepLabV3	93%	83%	89%
Proposed System	96%	89%	91%

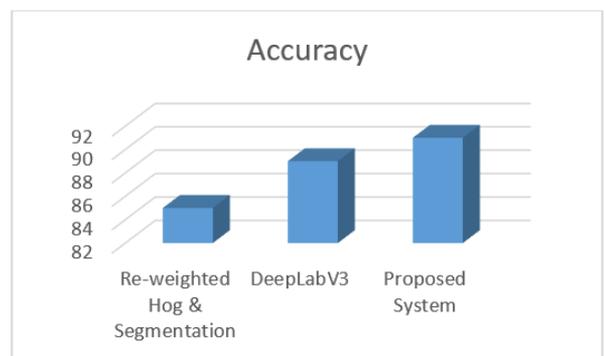


Figure 20. Performance analysis based on accuracy

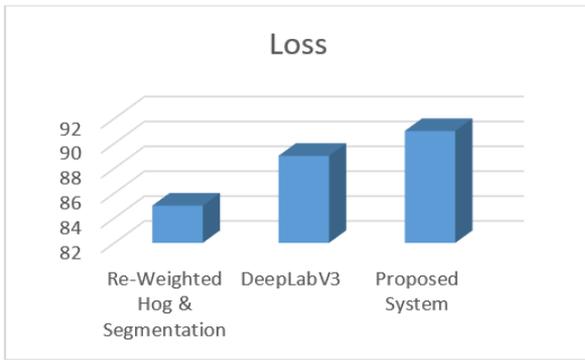


Figure 21. Performance evaluation based on loss

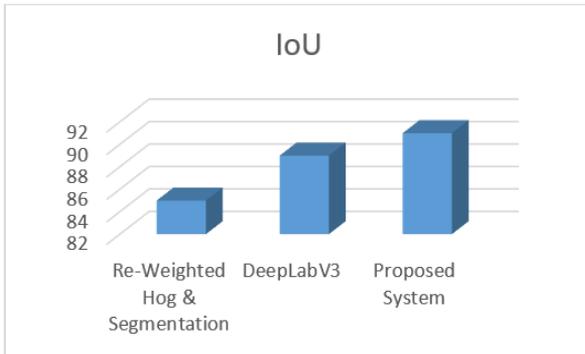


Figure 22. Analysis of Intersection over Union (IoU) of various data sets

The comparative analysis of the proposed system with other segmentation methods called Re-weighted HoG & segmentation method and DeeplabV3 methods are shown in Table 2. The performance metrics considered for the comparison are accuracy, precision and IoU. The Histogram of Oriented Gradient (HoG) is a core descriptor for object detection, which is re-weighted using the higher-level information getting from image segmentation. Whereas the DeeplabV3 is an architecture created for doing semantic segmentation, which is capable to segment the multiple scale images. The percentage of accuracy, precision and the IoU value of the former method is obtained as 82, 74 and 85 respectively and the same for the latter method is found as 93, 83 and 89. Comparing the proposed system with the above methods, the improvement in accuracy (96%), precision (89%) and IoU (91%) are obtained. The graphical representations of performance analysis of the above methods and the proposed system are shown in Figure 20, Figure 21 and Figure 22 respectively.

5. CONCLUSIONS

To provide an accurate and effective quality control in industries, a semantic segmentation technique using high calibre U-NET architecture is proposed for metal surface anomaly detection. The developed system has the highest degree of accuracy when predicting the existence of anomalies like dents, cracks, and scratches on surfaces. The high calibre U-NET model has been evaluated using different parameters like accuracy, loss and IoU. In order to assure the better performance, the proposed model architecture has been trained and tested using two publicly available data sets named NEU

data base and Kolektor data base for metal surfaces. According to the proposed system, the future directions of the research can be as follows: The segmented output from the proposed system can be further processed and utilized for the classification with the help of advanced deep learning models. Also, the architecture for semantic segmentation can be made in such a way to provide predictions without any loss in accuracy with the application of various data sets, where the captured images have variation in lighting condition as well as the background with complex features.

REFERENCES

- [1] Park, M., Jin, J.S., Au, S.L., Luo, S., Cui, Y.(2009). Automated defect inspection systems by pattern recognition. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 2(2): 31-42.
- [2] Tsa, D.M., Wu, S.K. (2000). Automated surface inspection using Gabor filters. *The International Journal of Advanced Manufacturing Technology*, 16(7): 474-482. <https://doi.org/10.1007/s001700070055>
- [3] Tsai, D.M., Huang, T.Y. (2003). Automated surface inspection for statistical textures. *Image and Vision computing*, 21(4): 307-323. [https://doi.org/10.1016/S0262-8856\(03\)00007-6](https://doi.org/10.1016/S0262-8856(03)00007-6)
- [4] Iivarinen, J. (2000). Surface defect detection with histogram-based texture features. In *Intelligent Robots and Computer Vision XIX: Algorithms, Techniques, and Active Vision*, 4197: 140-145. <https://doi.org/10.1117/12.403757>
- [5] Jie, L., Siwei, L., Qingyong, L., Hanqing, Z., Shengwei, R. (2009). Real-time rail head surface defect detection: A geometrical approach. In *2009 IEEE International Symposium on Industrial Electronics*, Seoul, pp. 769-774. <https://doi.org/10.1109/ISIE.2009.5214088>
- [6] Samarawickrama, Y.C., Wickramasinghe, C.D. (2017). Matlab based automated surface defect detection system for ceramic tiles using image processing. In *2017 6th National Conference on Technology and Management (NCTM) Malabe, Sri Lanka*, pp. 34-39. <https://doi.org/10.1109/NCTM.2017.7872824>
- [7] Schlegl, T., Seeböck, P., Waldstein, S.M., Schmidt-Erfurth, U., Langs, G. (2017). Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In *International conference on information processing in medical imaging* Boone, NC, USA, pp. 146-157. <https://doi.org/10.1007/978-3-319-59050-9>
- [8] Huang, Z., Xu, H., Su, S., Wang, T., Luo, Y., Zhao, X. (2020). A computer-aided diagnosis system for brain magnetic resonance imaging images using a novel differential feature neural network. *Computers in biology and medicine*, 121: 103818-103818. <https://doi.org/10.1016/j.compbiomed.2020.103818>
- [9] Wheeler, B.J., Karimi, H.A. (2020). Deep learning-enabled semantic inference of individual building damage magnitude from satellite images. *Algorithms*, 13(8): 195-195. <https://doi.org/10.3390/a13080195>
- [10] Zhang, J., Yang, X., Li, W., Zhang, S., Jia, Y. (2020). Automatic detection of moisture damages in asphalt pavements from GPR data with deep CNN and IRS method. *Automation in Construction*, 113: 103119-103119. <https://doi.org/10.1016/j.autcon.2020.103119>

- [11] Park, S.E., Eem, S.H., Jeon, H. (2020). Concrete crack detection and quantification using deep learning and structured light. *Construction and Building Materials*, 252, 119096-119096. <https://doi.org/10.1016/j.conbuildmat.2020.119096>
- [12] Cao, C., Ouyang, Q., Hou, J., Zhao, L. (2020). Visual Locating of Reactor in an Industrial Environment Using the Composite Method. *Sensors*, 20(2): 504-504. <https://doi.org/10.3390/s20020504>
- [13] Ünver, H.M., Ayan, E. (2019). Skin lesion segmentation in dermoscopic images with combination of YOLO and grabcut algorithm. *Diagnostics*, 9(3): 72. <https://doi.org/10.3390/diagnostics9030072>
- [14] Tao, T., Dong, D., Huang, S., Chen, W. (2020). Gap detection of switch machines in complex environment based on object detection and image processing. *Journal of Transportation Engineering, Part A: Systems*, 146(8): 04020083-04020083. <https://doi.org/10.1061/JTEPBS.0000406>
- [15] Zhang, H.W., Zhang, L.J., Li, P.F., Gu, D. (2018). Yarn-dyed fabric defect detection with YOLOV2 based on deep convolution neural networks. In 2018 IEEE 7th data driven control and learning systems conference (DDCLS) Enshi, China, pp. 170-174. <https://doi.org/10.1109/DDCLS.2018.8516094>
- [16] He, D., Xu, K., Zhou, P. (2019). Defect detection of hot rolled steels with a new object detection framework called classification priority network. *Computers & Industrial Engineering*, 128: 290-297. <https://doi.org/10.1016/j.cie.2018.12.043>
- [17] Mandal, V., Uong, L., Adu-Gyamfi, Y. (2018). Automated road crack detection using deep convolutional neural networks. In 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, pp. 5212-5215. <https://doi.org/10.1109/BigData.2018.8622327>
- [18] Dai, W., Mujeeb, A., Erdt, M., Sourin, A. (2020). Soldering defect detection in automatic optical inspection. *Advanced Engineering Informatics*, 43: 101004-101004. <https://doi.org/10.1016/j.aei.2019.101004>
- [19] Huang, Z., Li, F., Luan, X., Cai, Z. (2020). A Weakly Supervised Method for Mud Detection in Ores Based on Deep Active Learning. *Mathematical Problems in Engineering*. <https://doi.org/10.1155/2020/3510313>
- [20] Qiao, R., Ghodsi, A., Wu, H., Chang, Y., Wang, C. (2020). Simple weakly supervised deep learning pipeline for detecting individual red-attacked trees in VHR remote sensing images. *Remote Sensing Letters*, 11(7): 650-658. <https://doi.org/10.1080/2150704X.2020.1752410>
- [21] Majidifard, H., Jin, P., Adu-Gyamfi, Y., Buttlar, W.G. (2020). Pavement image datasets: A new benchmark dataset to classify and densify pavement distresses. *Transportation Research Record*, 2674(2): 328-339. <https://doi.org/10.1177/0361198120907283>
- [22] Jing, J., Zhuo, D., Zhang, H., Liang, Y., Zheng, M. (2020). Fabric defect detection using the improved YOLOv3 model. *Journal of engineered fibers and fabrics*, 15: 1558925020908268-1558925020908268. <https://doi.org/10.1177/1558925020908268>
- [23] He, Y., Zhou, Z., Tian, L., Liu, Y., Luo, X. (2020). Brown rice planthopper (*Nilaparvata lugens* Stal) detection based on deep learning. *Precision Agriculture*, 21(6): 1385-1402. <https://doi.org/10.1007/s11119-020-09726-2>
- [24] Véstias, M.P. (2019). A survey of convolutional neural networks on edge with reconfigurable computing. *Algorithms*, 12(8): 154-154. <https://doi.org/10.3390/a12080154>
- [25] Jeon, M., Jeong, Y.S. (2020). Compact and accurate scene text detector. *Applied Sciences*, 10(6): 2096-2096. <https://doi.org/10.3390/app10062096>
- [26] Pan, Y., Zhang, X., Cervone, G., Yang, L. (2018). Detection of asphalt pavement potholes and cracks based on the unmanned aerial vehicle multispectral imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(10): 3701-3712. <https://doi.org/10.1109/JSTARS.2018.2865528>
- [27] Ersoz, A.B., Pekcan, O., Teke, T. (2017). Crack identification for rigid pavements using unmanned aerial vehicles. In *IOP Conference Series: Materials Science and Engineering*, Prague, Czech Republic, 236(1): 012101-012101. <https://doi.org/10.1088/1757-899X/236/1/012101>
- [28] Dorafshan, S., Maguire, M., Qi, X. (2016). Automatic surface crack detection in concrete structures using OTSU thresholding and morphological operations, College of Engineering. <https://doi.org/10.13140/RG.2.2.34024.47363>
- [29] Dorafshan, S., Maguire, M., Hoffer, N.V., Coopmans, C. (2017). Fatigue crack detection using unmanned aerial systems in under-bridge inspection. College of Engineering.
- [30] Kim, H., Lee, J., Ahn, E., Cho, S., Shin, M., Sim, S.H. (2017). Concrete crack identification using a UAV incorporating hybrid image processing. *Sensors*, 17(9): 2052-2052. <https://doi.org/10.3390/s17092052>
- [31] Murao, S., Nomura, Y., Furuta, H., Kim, C.W. (2019). Concrete crack detection using UAV and deep learning. 13th International Conference on Applications of Statistics and Probability in Civil Engineering(ICASP13), Seoul, South Korea, pp. 1-8. <https://doi.org/10.22725/ICASP13.029>
- [32] Parker, L., Butterworth, J., Luo, S. (2019). Fly safe: Aerial swarm robotics using force field particle swarm optimisation. *arXiv preprint arXiv:1907.07647*. <http://arxiv.org/abs/1907.07647>
- [33] Kim, B., Cho, S. (2018). Automated vision-based detection of cracks on concrete surfaces using a deep learning technique. *Sensors*, 18(10): 3452. <https://doi.org/10.3390/s18103452>
- [34] Liu, H., Zhang, Y. (2019). Image-driven structural steel damage condition assessment method using deep learning algorithm. *Measurement*, 133: 168-181. <https://doi.org/10.1016/j.measurement.2018.09.081>
- [35] Dorafshan, S., Thomas, R.J., Maguire, M. (2018). Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete. *Construction and Building Materials*, 186: 1031-1045. <https://doi.org/10.1016/j.conbuildmat.2018.08.011>
- [36] Huang, H.W., Li, Q.T., Zhang, D.M. (2018). Deep learning based image recognition for crack and leakage defects of metro shield tunnel. *Tunnelling and underground space technology*, 77: 166-176. <https://doi.org/10.1016/j.tust.2018.04.002>
- [37] Gao, Y., Mosalam, K.M. (2018). Deep transfer learning

for image-based structural damage recognition. *Computer-Aided Civil and Infrastructure Engineering*, 33(9): 748-768. <https://doi.org/10.1111/mice.12363>

[38] Dung, C.V. (2019). Autonomous concrete crack

detection using deep fully convolutional neural network. *Automation in Construction*, 99: 52-58. <https://doi.org/10.1016/j.autcon.2018.11.028>