RESEARCH ARTICLE

Sequence of U-Shaped Convolutional Networks for Assessment of Degree of Delamination Around Scribe

Veronika Rozsivalova[1] · Petr Dolezel[1] · Dominik Stursa[1] · Pavel Rozsival[1]

Abstract

The application of protective layers is the primary method of keeping metallic structures resistant to degradation. The measurement of the layer resistance to delamination is one of the important indicators of the protection quality. Therefore, ISO 4628 standard has been issued to handle and quantify the main coating defects. Here, an innovative assessment of degree of delamination around a scribe according to ISO 4628 standard has been practically realized. It utilizes an computer-driven deep learning-based method. The assessment method is composed of two shallow U-shaped convolutional networks in a row; the first for preliminary and the second for refined detection of delamination area around a scribe. The experiments performed on 586 samples showed that the proposed sequence of U-shaped convolutional networks meets the edge computing standards, provides good generalization capability, and provides precise delamination area detection for a large variability of surfaces.

Keywords Coil coating · Delamination · Semantic segmentation · Deep learning

Abbreviations

EN	European Standard (from the German name Europäische Norm)
FCN	Fully convolutional neural network
FN	False negative
FP	False positive
ISO	International Organization for Standardization
RGB	Red, green, blue color model
TN	True negative
TP	True positive

✉ Petr Dolezel
petr.dolezel@upce.cz

Veronika Rozsivalova
veronika.rozsivalova@student.upce.cz

Dominik Stursa
dominik.stursa@upce.cz

Pavel Rozsival
pavel.rozsival@upce.cz

1   Faculty of Electrical Engineering and Informatics, University of Pardubice, Studentska 95, 532 10 Pardubice, Czech Republic

1 Introduction

Coil coating technology is a process of coating protective layers (e.g., primers, paints, epoxides, plastics, laminates, etc.) over metallic substrates, see [12, 24] for more information. The process of coating usually results in a multilayered structure consisting of a primer and top coat (or several top coat layers). This creates very beneficial effects although the coating thickness is very low, reaching tens of microns (about 30-40 μm), as it has been argued in [22]. The technology of coil coating is a very effective approach to improve the corrosion resistance of metals and prolong their lifetime. It is widely used in roofing systems, architectural elements, and interior applications, since it provides great flexibility, low added weight to a coated element, and various possibilities of special effects (color shifting, shining, and sparkling). Coil coated materials can also be easily bent, twisted, or formed without damaging their surface.

Nevertheless, the protective layer can be mechanically damaged by scribes or scratches and then the corrosion caused by water, sun, salt or corrosive gases and vapors can make irreversible changes to the metallic structure, see [3]. The resulting defects can be very diverse, and range from chalking and blistering to flaking or rusting of the coated metal. Hence, it is important to test coated surfaces under the conditions which closely simulate outdoor exposition.

The degradation resistance of coil coated metallic structures is evaluated according to the European Standard EN 13523-8 "Coil coated metals. Test methods. Resistance to salt spray (fog)" by exposing a test specimen to a salt fog at a defined temperature for a given time. During this procedure, a specimen is tested using the International Organization for Standardization ISO 4628 standard "Paints and varnishes. Evaluation of degradation of coatings. Designation of the quantity and size of the defects and intensity of uniform changes to appearance". The ISO 4628 standard has been defined for the purpose of the assessment and quantification of the main defects that can occur in coatings.

The ISO 4628 standard is a comprehensive document, which offers a more than necessary number of approaches for evaluation of every possible type of coating defect. In most cases, a manual evaluation of samples is recommended, e.g., in [3, 9, 22]. Unexpectedly, a limited number of techniques for an automated assessment described in the ISO 4628 standard have yet to be practically realized. For example, [16] utilized a non-contact laser profilometer to measure the profile differences in the sample and, consequently, the relief defects in the coating. The profilometer output was processed using commercial software for the image processing (Adobe Pro Suite, Adobe inc.). In [5], the authors used an Epson Perfection 3170 scanner for RGB (Red, Green, Blue Color Model) image acquisition and the corrosion in the images was evaluated using Adobe Photoshop Elements 2.0 and an open source software ImageJ. Classical approaches of the image processing for degradation evaluation were also used in works [7, 14]. Additionally, [25] presented an approach to quantify corrosion susceptibility based upon electrical resistance and colorimetric outputs.

A common drawback of the above-mentioned approaches is the emphasis on a limited range of materials, surfaces, or colors of the coating. Conversely, industrial entities ask for a generally applicable robust system. In accordance with this trend, deep neural networks have proven themselves to be very successful in image processing tasks dealing with a wide scope of input data. In response to this corrosion detection demand, [20] presented a semantic segmentation deep learning-based approach and an efficient image labeling tool to detect, segment, and evaluate corrosion in the images.

Their method was proposed for classifying each pixel of a corrosion segment into user-prescribed categories such as light corrosion, medium corrosion, and heavy corrosion. The authors in [19] introduced an automated detection of corrosion in used nuclear fuel storage canisters, by employing residual neural networks. Furthermore, a method for inspection and characterization of external corrosion in pipelines using a deep learning-based method is presented in [2].

However, to the best of our knowledge, a deep learning-based method for automatic or autonomous evaluation of degradation of coatings according to the ISO 4628 standard has not yet been introduced. Therefore, the aim of this article is to present an autonomous efficient deep learning-based method, which implements the ISO 4628-8 standard "Assessment of degree of delamination and corrosion around a scribe or other artificial defect". This standard is a part of the ISO 4628 standard, and it specifies a method for assessing delamination and corrosion around a scribe on a coated panel, or other coated test specimens, caused by a corrosive environment. The specification includes the criteria for evaluating the intensity and quantity of the delamination of the coating or corrosion of the specimen. A numerical scale from 0 to 5 (represented in millimeters) is adopted for evaluation, where 0 stands for the absence of changes, while 5 means defects so notable that further discrimination is not reasonable. A general pipeline of the presented method is depicted in Fig. 1. As indicated in the figure, the presented method requires a metal test specimen inserted into the RGB image acquisition device. Subsequently, the RGB image of the specimen is processed by a deep learning device, which performs image semantic segmentation. A black and white image is a result of this operation, where white color represents the area of delamination. Finally, the black and white image is processed by an evaluation device, which grades the specimen according to the white area ratio in the image.

The presented method is applicable to a wide range of coatings, regardless of color, asperity, or reflectivity. Additionally, the method requires no special hardware for data acquisition and is computationally efficient enough to be implemented using the edge computing tools. Examples of the various types of coatings, that are considered within the presented method, are shown in Fig. 2.
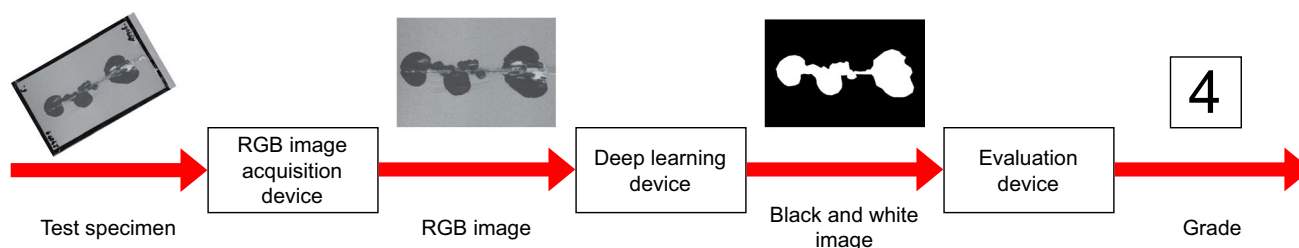
Fig. 1 A pipeline of the presented autonomous method for the assessment of the degree of delamination around a scribe

Fig. 2 Examples of test specimens, that are considered within the presented deep learning-based method
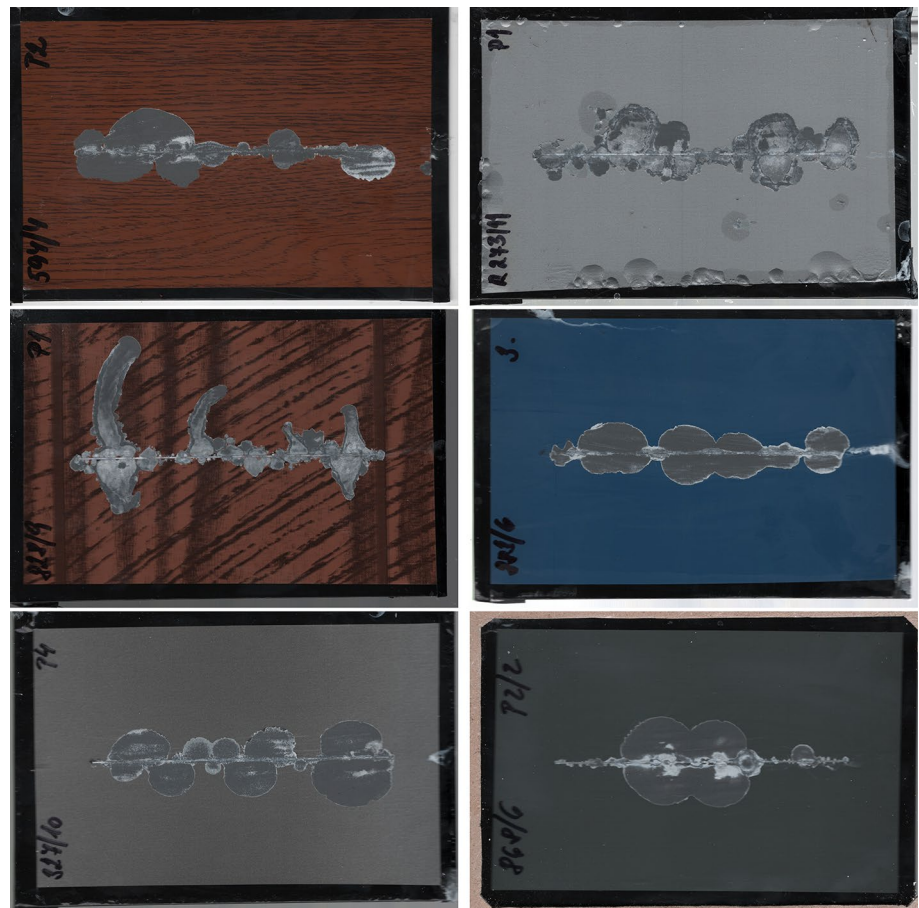
2 Related Work

Image semantic segmentation is an essential procedure in the presented method. The objective of semantic segmentation is generally to convert the original image into a representation, which is easier to analyze. Specifically, each pixel of the original image is labeled with a corresponding class. Hence, by densely labeling all the pixels of an image, abstract representations of objects or their shapes in the original image can be constructed.

As explained by [17], deep learning-based methods to image semantic segmentation can be divided into tens of groups, according to many distinctive criteria. The selected deep learning architectures, which have already proven to be effective with various practical tasks, are provided in the following paragraphs.

One of the first deep learning-based approaches for image semantic segmentation was the Fully Convolutional Network (FCN) proposed in [15]. The authors replaced existing convolutional neural network architectures by fully connected layers with convolutional layers. Their topology generated a spatial segmentation map for the original image instead of classification scores as the result.

At the present time, most convolutional neural network models for image semantic segmentation implement a convolutional encoder-decoder procedure. The SegNet model, proposed in [1], consists of an encoder part, which is topologically identical to the thirteen convolutional layers in the VGG16 network, see [23], and a corresponding decoder part followed by a pixel-wise classification layer. The similar procedure was implemented with the ResNet model introduced in [10], MobileNet model introduced in [11], or DenseNet model introduced in [13] as a backbone. As an important part of encoder-decoder family of methods, U-shaped models attract remarkable attention. These models enhance their structure into the U-shaped topology with deconvolutional or up-sampling layers. The U-Net, designed by [21], presented also the so-called skip connection network topology, and the BiSeNet [26] brought two parallel paths in its topology - a spatial path was designed to cover the spatial information and generate high-resolution features, while a context path was implemented to obtain a necessary receptive field. Besides, various versions of attention mechanism were successfully implemented into semantic segmentation problems, especially in cases where the same objects occurred in different scales, with different contrast, or in different context. For example, authors in [18] introduced

an attention mechanism to an existing U-Net architecture and showed a substantial rise of the prediction performance in comparison to the original U-Net. Last of all, bearing in mind edge computing phenomenon, some authors tried to reduce the number of parameters of the existing architectures while keeping the performance. For example, [4] presented the Squeeze U-Net, which reduced the number of parameters from 30 million (original U-Net) to 2.59 million with similar accuracy.

3 Materials and Methods

3.1 Standard ISO 4628-8

Various testing procedures of coil coated metals include, among others, measurement of degradation resistance to salt fog (standard EN 13523-8). Here, one of the typically evaluated quantities is the assessment of the degree of delamination around a scribe, which is defined by the ISO 4628-8 standard. The test specimen is horizontally scribed with a sharp edge and exposed to a corrosive environment of salt fog. After a defined period of time, the specimen is cleaned with tap water, and the residues of water are dried using compressed air. A loose coating is removed with a blade held at a sharp angle. See Fig. 2 for an example of specimens prepared by this procedure.

As stated by the ISO 4628-8 standard, two variants of the assessment of degree of delamination around a scribe can be implemented.

3.1.1 First Variant

The width of the area of delamination has to be measured at a minimum of six points uniformly distributed along the scribe. Subsequently, the arithmetic mean is determined and the resulting value is designated as the mean overall width of the zone of delamination, d_1, in millimeters.

The degree of delamination d, in millimeters, can be calculated using the equation

$$d = \text{round}\left(\frac{d_1 - w}{2}\right), \tag{1}$$

where w is the width of the original scribe, in millimeters. See Fig. 3 for an example of application.

3.1.2 Second Variant

The area of delamination is determined in this approach. The standard proposes laying a transparent millimeter-grid paper over the plate and counting the number of squares
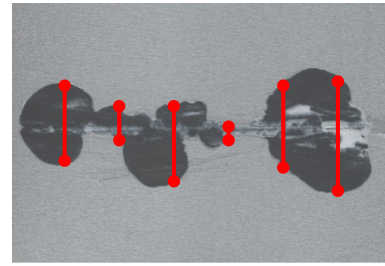
Fig. 3 The mean overall width of the zone of delamination is defined as a sum of the abscissa lengths shown in the figure divided by six

Fig. 4 White area is the area of delamination, the scribe is marked by red

corresponding to the delamination area. Then, the degree of delamination d, in millimeters, can be calculated using

$$d = \text{round}\left(\frac{A_d - A_l}{2l}\right), \tag{2}$$

where A_d is the area of delamination, including the area of the scribe (in square millimeters), A_l is the area of the scribe in the evaluated area in square millimeters, and l is the length of the scribe in the evaluated area (in millimeters). See Fig. 4 for an example of application.

3.2 Dataset

To design a deep learning-based method for evaluation of coating degradation it is necessary to acquire an extensive dataset of various annotated specimens with diversified color, asperity, or reflectivity of coatings and dispersed degrees of delamination. To do so, 586 coated specimens of 150×100 mm were prepared. The samples contained white, blue, green, gray, brown, dark blue, orange, red, black and yellow coatings in fine and coarse (structured) variants. A narrow horizontal scribe (0.5 mm width) was made with an iron nail through the coating of every sample to expose the uncoated material. The specimens were then inserted into a salt fog chamber for a defined period of time—120 h, 240 h, 480 h, 720 h and 1440 h. After exposure, the specimens were cleaned as described above, scanned by a flatbed scanner and

annotated using a custom-made application named Delamination Labeler (see Fig. 5 for an example). As a result, 586 input-output pairs were created. Finally, the dataset was randomly divided into a training set (456 samples) and testing set (130 samples).

3.3 Sequence of Convolutional Neural Networks for Delamination Area Detection

To autonomously address the problem of image semantic segmentation, as shown in Fig. 1, a sequence of deep learning networks is proposed. The sequence consists of two consecutive neural networks: the Raw Module and the Refinement Module. The Raw Module provides the preliminary segmentation results of the input image. Subsequently, the input to the Refinement Module is attained by the concatenation of the input image and the output from the Raw Module. The Refinement Module is expected to improve the segmentation accuracy by correcting the segmentation result of the Raw Module according to the information from the input image. See Fig. 6 for a brief sequence of the proposed method.

The structure of both modules is inspired by the U-Net. In comparison to the original U-Net, a set of changes to the original U-Net topology is proposed to improve the inference time and to lower the memory requirements. As a result, each module includes less than 2 million parameters contrary to 30 million in the original U-Net.

3.3.1 Raw Module

Although the original U-Net created by [21] is an indisputable reference for the architecture of the Raw Module, there are several modifications implemented to prepare the architecture for edge computing applications. Overall, the Raw Module is shallower and contains fewer filters.

A 288×288 px RGB image of the test specimen is considered as an input to the module. The encoder part of the module uses a series of convolution operations (kernel size $= 3 \times 3$, rectified linear unit used as an activation function). A max-pooling layer (pool size $= 2 \times 2$) is always included after two convolutional layers. The starting number of filters is 32 in the first convolutional layer, and this number is doubled after each max-pooling layer. Altogether, three max-pooling layers are implemented in the encoder part of the module. Additionally, three dropout layers (dropout rate $= 0.2$) are included for training purposes to lower the possibility of overfitting.

The decoder part of the module always starts with the concatenation block, which connects the output signal from the previous layer with the skip connection from the encoder part of the module. Then, two convolutional layers follow, and the signal is up-sampled to 2×2 using a bilinear interpolation. After every up-sampling operation, the number of filters in the convolutional layers is halved. Again, some dropout layers are included for training purposes. As a last operation, the signal from the decoder is activated using the hyperbolic tangent activation function.

The detailed scheme of the topology is shown in Fig. 7.

3.3.2 Refinement Module

The obvious idea for this module is to implement the same architecture as used in the Raw Module. Only the input would be a concatenation of the input image and the output from the Raw Module. This approach will be referred to

Fig. 5 The Delamination Labeler - a custom-made application for data annotation
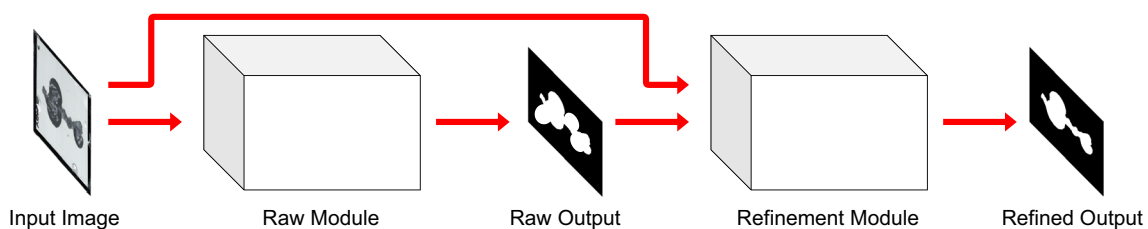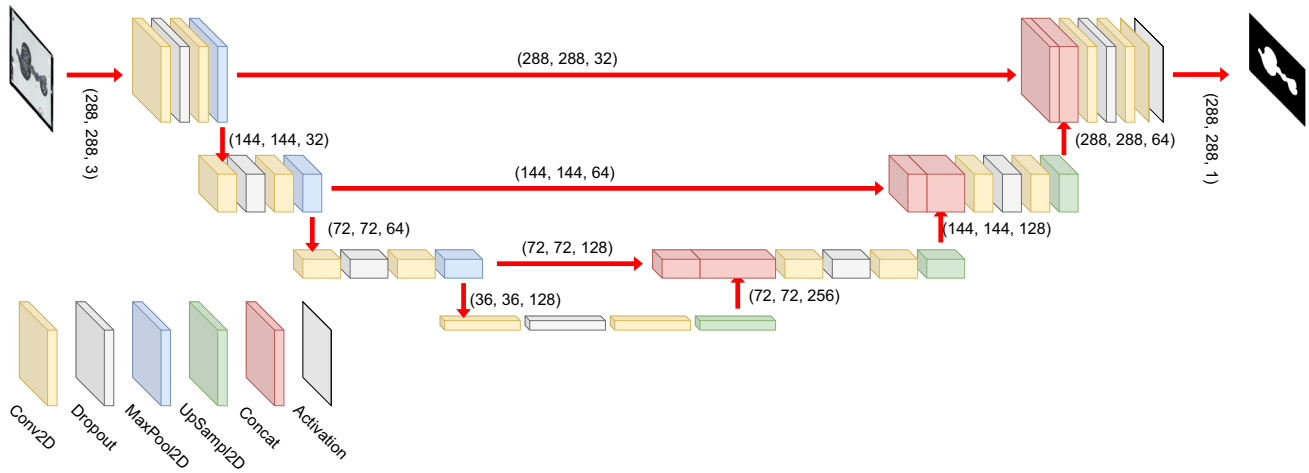

Fig. 6 The proposed sequence of U-shaped convolutional networks

Fig. 7 The Raw Module. Numbers in the brackets refer to the dimensions of the signal

as the Full-Size Refinement Module in the following text. Besides, another approach to the Refinement Module architecture is also proposed. Since the degree of delamination according to the ISO 4628 standard is provided as an integer number from a numerical scale between 0-5 millimeters, the authors believe that the resolution reduction at the output of the Refinement Module may not significantly reduce the accuracy of the provided degree of delamination. On the other hand, it may considerably decrease the inference time of the Refinement Module and the computational complexity of the operations performed in the Evaluation device (see Fig. 1).

Therefore, in comparison to the Raw Module, the up-sampling operations in the decoder part are omitted, except for the first one. On the other hand, a max-pooling layer (pool size = 4 × 4) is added to the first skip connection. A

max-pooling layer (pool size = 2 × 2) is also added to the second skip connection to match the signal dimensions in the decoder part of the module. This module (see Fig. 8) will be referred to as the Reduced Refinement Module in the following text.

3.4 Neural Network Training

The training of both modules is performed from scratch, minimizing a binary cross entropy loss function. As an initialization method, a normal distribution function with mean set to 0 and standard deviation set to 0.05 is applied. The Adam optimizer is implemented for the training. An initial learning rate of 0.001 is selected, and an exponential decay rate for the first and second moment estimates is set to 0.9 and 0.999, respectively.
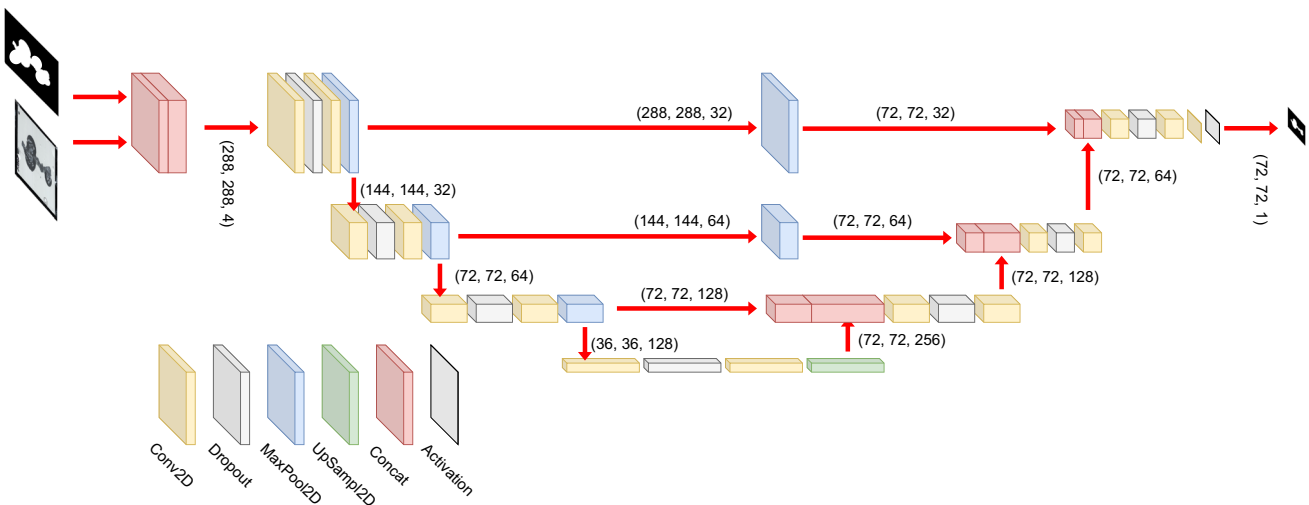
Fig. 8 The Reduced Refinement Module. Numbers in the brackets refer to the dimensions of the signal

Data augmentation is used to avoid overfitting during training. Specifically, random rotation (range of a rotation angle $= \pm 5°$), horizontal and vertical flipping, horizontal and vertical translation (up to $\pm 5\%$ of the input image height and width, respectively), and horizontal and vertical shear (shear intensity $= 0.1$) are applied.

Firstly, the Raw Module was trained using the training set. A random 15 % training sample ratio were used as the validation set and the rest was applied to 800 epochs of training (batch size $= 8$). After every epoch, the loss function over the validation set, defined as binary cross entropy, was evaluated. At the end, the model instance (i.e. the values of weights and biases) with the best performance value over the validation set during the whole training session was accepted for further processing. Moreover, 10 repetitions of the training session were carried out to deal with the stochastic behavior of the training process.

Secondly, the responses of the Raw Module over the training set were evaluated and the results were stored to prepare necessary input samples for the training of the Refinement Module. Each output had to be concatenated with the original input image to get an expected array of the shape (288, 288, 4). Subsequently, the Refinement Module was trained with these input samples under the same conditions as above.

The performance of the proposed method was evaluated on a personal computer with Intel Core i5-8600K (3.6 GHz) CPU, internal memory 16 GB DDR4 330 (2666 MHz), video card NVIDIA PNY Quadro P5000 16 GB GDDR5 PCIe 3.0. For evaluation of the inference time, NVIDIA Jetson NANO computer with Quad-core ARM A57 1.43 GHz CPU and 4 GB RAM was used. This single board computer is a generally accepted hardware for the benchmarking of deep learning-based edge computing applications.

3.5 Competitive Approach

To prove the proposed sequence of neural networks for delamination area classification, the performance is compared with two state-of-the-art architectures - SegNet introduced by [1] and U-Net firstly used in [21]. These networks are trained under the same conditions as the Raw Module listed above.

In addition, two classical image processing approaches, which have already been applied to the automated approach to assess the degree of delamination around a scribe [8], are included for comparison. The first approach implements an active contours method without edges proposed by [6]. Adversely, the other approach proposed in the cited reference is based more on "brutal force". It implements a set of very basic image processing operations, including subtracting the image of the sample from the image acquired at the beginning of the degradation process, filtration, thresholding

and some other image adjustment operations. In accordance with the information provided in [8], these methods will be referred to as First Approach and Second Approach, respectively.

3.6 Evaluation Metrics

Two key aspects of the presented image semantic segmentation method are classification performance and time-complexity. A common practice for the evaluation of the classification performance is calculation of accuracy over a testing set (a dataset independent of the training set). Semantic segmentation is basically a pixelwise classification of the image content. For the classification of the delamination area, a true positive pixel is a pixel annotated as true in the target sample as well as in the refined output. A false positive pixel is annotated as true in the refined output, but is false in the target sample. A true negative pixel is a pixel annotated as false in the target sample as well as in the refined output. Finally, a false negative pixel is annotated as false in the refined output, but is true in the target sample. Then, the accuracy is given as

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}, \tag{3}$$

where TP is the number of true positive pixels in the tested sample, FN is the number of false positive pixels in the tested sample, FP is the number of false positive pixels in the tested sample, and TN is the number of true negative pixels in the tested sample.

To evaluate the classification performance comprehensively, additional measures are also considered:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{4}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{5}$$

$$\text{F1-score} = \frac{2}{\text{Recall}^{-1} + \text{Precision}^{-1}}. \tag{6}$$

To evaluate time-complexity of the presented image semantic segmentation approach to the competitive methods, a relative inference time is given as

$$\tau = \frac{t_{\text{C}}}{t}, \tag{7}$$

where t_{C} is the total inference time of a selected competitive approach over the testing set, and t is the total inference time of the presented approach with the Reduced Refinement Module.

3.7 Pipeline for the Autonomous Method for the Assessment of Degree of Delamination around a Scribe

The whole pipeline of the autonomous method for the assessment of degree of delamination around a scribe can be outlined considering the following information (as proposed in Fig. 1).

As a first step, an RGB image acquisition device must be carried out. Since no special requirements are demanded, any generic scanner or RGB camera can be considered. An interesting possibility would be a mobile phone with an integrated camera, since the whole method could be then deployed in a single body of common hardware.

Then, the image semantic segmentation approach needs to be implemented in the deep learning device. The presented sequence of the Raw Module and the Refinement Module can be applied using a wide range of computers from custom single board solutions to robust cloud-computing services.

As a last step, the evaluation device is expected to process the segmented image into a specific value of the degree of delamination. This can be achieved by the generic implementation of either Eq. (1) or Eq. (2). From the practical point of view, the possibility, which is implemented by Eq. (2), is probably less accurate, but it easily enables the manual correction of the potentially inaccurate result. This situation is depicted in Fig. 9.
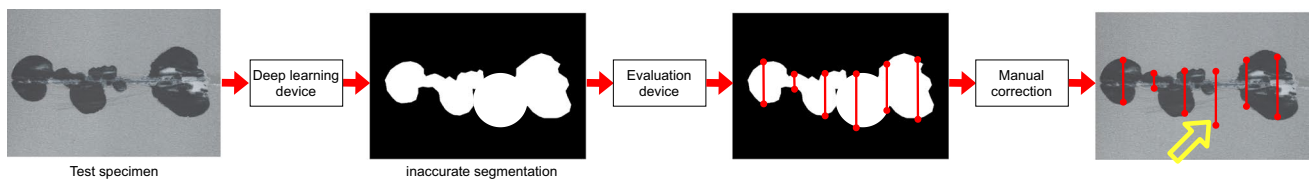
4 Results and Discussion

The proposed sequence of the convolutional neural networks for the detection of delamination area, as well as the competitive approaches, were trained according to the procedure addressed in Sect. 3.4 with the dataset described in Sect. 3.2. To show the capability of the proposed approach, the resulting values of the evaluation measures described in Sect. 3.6 are summarized in Table 1. Note that the relative inference time is not evaluated for the last two rows of the table, since these methods are not based on deep learning and the implementation on the hardware developed for the deep learning-based applications (see Sect. 3.4) would be misleading.

Furthermore, absolute frequencies of differences between the ground truth degree of delamination and degrees provided by each considered method using Eq. (2) were demonstrated in Table 2. Specifically, the degree of delamination for each sample was assessed manually using the ISO 4628-8 standard, and then using the presented approach. The resulting values, in millimeters, were subtracted to demonstrate the difference between the ground truth result and the result provided by the proposed method. Additionally, the sum of absolute differences $\Sigma|E|$ was determined to show the overall performance of the methods over the testing set. These values in Table 2 indicate the practical impact of the presented approach, since the clear number of errors, including their magnitudes, is explicitly depicted.

Some examples of the delamination detection process are shown in Fig. 10.

Fig. 9 The possibility of manual correction of the result—if the process is supervised by a human operator, the length of the incorrectly placed abscissa (yellow arrow) can be corrected manually

Table 1 Evaluation results

Segmentation method	Adapted from	mAcc	mPrec	mRec	mF1-sc	τ
Raw module & Full-size refinement module	Proposed	0.9906	0.8813	0.8835	0.8824	1.1972
Raw module & Reduced refinement module	Proposed	0.9892	0.8782	0.8747	0.8764	1.0000
U-Net	[21]	0.9893	0.8906	0.8777	0.8841	1.3310
SegNet	[1]	0.9755	0.7602	0.5452	0.6350	4.1056
First approach	[6, 8]	0.8929	0.7124	0.7880	0.7483	–
Second approach	[8]	0.9426	0.8563	0.5717	0.6856	–

Mean accuracies (mAcc), mean precisions (mPrec), mean recalls (mRec), mean F1-scores (mF1-sc) and relative inference time (τ) evaluated over the testing set

Table 2 Absolute frequencies of differences between ground truth degree of delamination and degrees provided by each considered method using Eq. (2)

	<− 2 mm	− 2 mm	− 1 mm	0 mm	1 mm	2 mm	>2 mm	ΣE
Raw module & Full-size refinement module	1	2	10	104	13	0	0	30
Raw module & Reduced refinement module	1	3	8	105	12	1	0	31
U-Net	1	4	5	109	3	2	6	41
SegNet	1	0	1	70	33	8	17	104
First approach	34	0	2	76	14	3	1	127
Second approach	3	0	0	72	40	10	5	84

ΣE means the sum of absolute errors provided by each method over the testing set

The evaluation results presented in Tables 1 and 2 speak in favor of the proposed approaches, together with U-Net. Using the considered testing dataset, there are only small differences between the results provided by the Raw Module and Full-size Refinement Module, Raw Module and Reduced Refinement Module, and U-Net.

Considering mean accuracy as the first metric, the Raw Module and Full-size Refinement Module provides a value higher than 0.99, while the other two aforementioned approaches have values just less than 0.99. However, note that all considered approaches provide far higher accuracy results in comparison to precision and recall. This phenomenon is caused by a naturally imbalanced dataset, since the area of delamination is significantly smaller than the rest of the image. Therefore, the number of true negative pixels is much greater than the number of positive pixels. In such situations, it is important to consider precision, recall, and their harmonic mean - F1 score. The best value of precision is provided by U-Net, followed closely by both variants of the proposed approach. Looking at recall, the Raw Module and Full-size Refinement Module delivers the best result. On the other hand, the recall results provided by SegNet and the Second Approach adopted from [8] are very low. Moreover, it is also necessary to consider the ratio of precision and recall values. In this case, both of the proposed methods deliver very beneficial results, i.e. the number of false positive pixels is in both cases very similar to the number of false negative pixels (if the number of false positive pixels is equal to the number of false negative pixels, the area of delamination is identified flawlessly). This feature clearly indicates the suitability of the proposed methods for delamination area detection.

Bearing in mind the relative inference time, the best result is naturally provided by the Raw Module and Reduced Refinement Module, followed by the Raw Module and Full-size Refinement Module. However, U-Net provides only 33 % worse inference time, which is far more acceptable in comparison to SegNet.

Looking at Table 2, in accordance with the aforementioned findings, the most suitable results are provided by the proposed approaches, in conjunction with U-Net. Apart from the low number of overall errors, the balanced distribution of positive and negative errors can be observed in the first three rows of the table. However, the sum of absolute errors $\Sigma|E|$ clearly distinguishes the proposed methods contrary to U-Net. In other words, if the solution using U-Net evaluates the degree of delamination incorrectly, this error is usually larger than when applying the presented methods. Apparently, SegNet and the methods adopted from [8] give significantly less acceptable results.

As a summary, an innovative assessment of degree of delamination around a scribe according to the ISO 4628 standard is practically realized in this study. This represents, to the authors' knowledge, the first comprehensive application of convolutional neural networks to this problem. The proposed methods for delamination area detection (the Raw Module and Full-size Refinement Module, Raw Module and Reduced Refinement Module) provide comparable or better results against the current state-of-the-art techniques. Additionally, the proposed methods are computationally more efficient, and less memory intensive compared to the well-known U-Net and SegNet.

When comparing the two presented methods, it is necessary to stress, that the Reduced Refinement Module produces only a negligible reduction in quality compared to the Full-size Refinement Module. However, it is less computationally and memory intensive, and additionally, it provides a $16\times$ smaller output image. This fact brings the potential of even more computational cost reductions in subsequent data processing [e.g., implementation of Eq. (2)].

Original input images (288 × 288 px).

Delamination detection using the Raw Module & Full-size Refinement Module (288 × 288 px).

Delamination detection using the Raw Module & Reduced Refinement Module (72 × 72 px).

Delamination detection using U-Net (288 × 288 px).

Delamination detection using SegNet (288 × 288 px).

Delamination detection using the First Approach (288 × 288 px) Dolezel et al (2018); Chan and Vese (2001).

Delamination detection using the Second Approach (288 × 288 px) Dolezel et al (2018).

Fig. 10 Examples of delamination detection using selected approaches

5 Conclusions

The presented results clearly showed the suitability of deep learning-based methods for the assessment of the degree of delamination around a scribe. All the metric values considered in this article indicate that the U-shaped convolutional neural networks are especially convenient for the fast and precise detection of the area of delamination of the coated surfaces. Additionally, the utilization of the sequence of two shallow U-shaped modules resulted in the delamination detection method, which meets the edge computing standards, provides good generalization capability, and shows good ratio of precision and recall. The method successfully provides precise delamination area detection for a large variability of surfaces.

Features of the presented deep learning-based delamination area detection method predetermine its implementation in affordable compact devices, which will integrate all main parts (RGB image acquisition device, deep learning device, evaluation device) into one solid body. Such an instrument can be easily used in both the industrial praxis and laboratory experiments taking place under special operational conditions (explosion hazard, dusty, humid or corrosive environment).

Data Availability The data presented in this study are available on request from the corresponding author. The data are not publicly available.

Declarations

Conflict of interest All authors declare that they have no conflicts of interest.

References

1. Badrinarayanan, V., Kendall, A., Cipolla, R.: Segnet: a deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **39**(12), 2481–2495 (2017). https://doi.org/10.1109/TPAMI.2016.2644615

2. Bastian, B., Ranjith, S., et al.: Visual inspection and characterization of external corrosion in pipelines using deep neural network. NDT E Int. ,(2019). https://doi.org/10.1016/j.ndteint.2019.102134

3. Bastos, A., Simões, A.: Effect of deep drawing on the performance of coil-coatings assessed by electrochemical techniques. Prog. Org. Coat. **65**(2), 295–303 (2009). https://doi.org/10.1016/j.porgcoat.2009.01.002

4. Beheshti, N., Johnsson, L.: Squeeze u-net: a memory and energy efficient image segmentation network. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp 1495–1504. https://doi.org/10.1109/CVPRW50498.2020.00190 (2020)

5. Blanchard, P., Hill, D., Bretz, G., et al.: Evaluation of corrosion protection methods for magnesium alloys in automotive applications, vol 9781118858943. https://doi.org/10.1002/9781118859803.ch94 (2014)

6. Chan, T.F., Vese, L.A.: Active contours without edges. IEEE Trans. Image Process. **10**(2), 266–277 (2001). https://doi.org/10.1109/83.902291

7. Cringasu, E.C., Dragomirescu, A., Safta, C.A.: Image processing approach for estimating the degree of surface degradation by corrosion. In: 2017 International Conference on Energy and Environment (CIEM), pp 275–278 (2017). https://doi.org/10.1109/CIEM.2017.8120791

8. Dolezel, P., Rozsival, P., Rozsivalova, V., et al.: Automatized approach to assessment of degree of delamination around a scribe. Adv. Intell. Syst. Comput. **868**, 227–236 (2018). https://doi.org/10.1007/978-3-030-01054-6_16

9. Hanus, J.: Selection and evaluation of singlelayer coating compositions in corrosive environments. Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis 59(5):53–64 (2011). https://doi.org/10.11118/actaun201159050053

10. He, K., Zhang, X., Ren, S., et al.: Deep residual learning for image recognition, pp. 770–778 (2016). https://doi.org/10.1109/CVPR.2016.90

11. Howard, A.G., Zhu, M., Chen, B., et al.: Mobilenets: efficient convolutional neural networks for mobile vision applications (2017). arxiv:1704.04861

12. Jandel, A.S.: Sustainability in the coil coatings sector. JOT Journal fuer Oberflaechentechnik **59**(8), 10–13 (2019). https://doi.org/10.1007/s35144-019-0274-3

13. Jégou, S., Drozdzal, M., Vazquez, D., et al.: The one hundred layers tiramisu: fully convolutional densenets for semantic segmentation. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp 1175–1183 (2017). https://doi.org/10.1109/CVPRW.2017.156

14. Kapsalas, P., Zervakis, M., Maravelaki-Kalaitzaki, P.: Evaluation of image segmentation approaches for non-destructive detection and quantification of corrosion damage on stonework. Corros. Sci. **49**(12), 4415–4442 (2007). https://doi.org/10.1016/j.corsci.2007.03.049

15. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 3431–3440 (2015). https://doi.org/10.1109/CVPR.2015.7298965

16. Maples, A., Williams, E., Rawlins, J.: Understanding scribe profile and tool type effects on visual corrosion assessments. J. Fail. Anal. Prev. **18**(6), 1401–1410 (2018). https://doi.org/10.1007/s11668-018-0528-z

17. Minaee, S., Boykov, Y., Porikli, F., et al.: Image segmentation using deep learning: a survey. IEEE Trans. Pattern Anal. Mach. Intell. (2021). https://doi.org/10.1109/TPAMI.2021.3059968

18. Oktay, O., Schlemper, J., Folgoc, L.L., et al.: Attention u-net: learning where to look for the pancreas (2018). arxiv:1804.03999

19. Papamarkou, T., Guy, H., Kroencke, B., et al.: Automated detection of corrosion in used nuclear fuel dry storage canisters using residual neural networks. Nucl. Eng. Technol. **53**(2), 657–665 (2021). https://doi.org/10.1016/j.net.2020.07.020

20. Rahman, A., Wu, Z., Kalfarisi, R.: Semantic deep learning integrated with RGB feature-based rule optimization for facility surface corrosion detection and evaluation. J. Comput. Civ. Eng. (2021). https://doi.org/10.1061/(ASCE)CP.1943-5487.0000982

21. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. Lecture Notes Comput. Sci. **9351**, 234–241 (2015). https://doi.org/10.1007/978-3-319-24574-4_28

22. Rossi, S., Deflorian, F., Fiorenza, J.: Environmental influences on the abrasion resistance of a coil coating system. Surf. Coat. Technol. **201**(16–17), 7416–7424 (2007). https://doi.org/10.1016/j.surfcoat.2007.02.008

23. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition (2015)

24. The National Coil Coating Association: NCCA promotes growth of coil coating industry. Light Metal Age **78**(5), 66–67 (2020)

25. Whitfield, M., Bono, D., Wei, L., et al.: High-throughput corrosion quantification in varied microenvironments. Corros. Sci. **88**, 481–486 (2014). https://doi.org/10.1016/j.corsci.2014.07.045

26. Yu, C., Wang, J., Peng, C., et al.: Bisenet: bilateral segmentation network for real-time semantic segmentation. Lecture Notes Comput. Sci. **11217**, LNCS:334–349 (2018). https://doi.org/10.1007/978-3-030-01261-8_20

Springer