Identification and localization of pitting corrosion on metallic surface using deep learning

N. V. Krysko, *Cand. Eng., Associate Prof., Dept. MT-7 "Welding Tehnology and Diagnostics"* [1], *e-mail: kryskonv@bmstu.ru;*
N. A. Shchipakov, *Cand. Eng., Associate Prof., Dept. MT-7 "Welding Tehnology and Diagnostics"* [1], *e-mail: shchipak@bmstu.ru;*
D. M. Kozlov, *Cand. Eng., Engineer* [1], *e-mail: kozlovdm@bmstu.ru;*
A. G. Kusyy, *Engineer* [1], *e-mail: kusyy@bmstu.ru*

[1] *Bauman Moscow State Technical University (Moscow, Russia)*

In this work, a computer vision system is proposed, which allows the identification and localization of pitting corrosion on metallic surface of the gas pipelines made of low carbon and low alloy steels. For this purpose, a dataset of 5,760 images of pipeline surface with and without pitting corrosion was collected. The developed convolutional neural network (CNN) architecture was trained on this dataset. The hyperparameters of this architecture were optimized using Bayesian optimization. The developed and optimized CNN architecture does not have a large number of trainable parameters compared to existing CNN-based architectures. Also, the developed architecture showed significantly higher accuracy of 98.44 %, when classifying images into images without corrosion and with pitting corrosion. The developed CNN outperformed most existing classifiers in its parameters. A pitting corrosion localization system was also developed using the "sliding windows" and "image pyramid" methods, which made it possible to localize areas with identified pitting corrosion on the surface of pipelines made of low carbon and low alloy steels, using the developed CNN without additional labeling of the data set. The proposed deep learning approach will eliminate the need for the operator to visually inspect the pipeline for pitting corrosion, which is costly and time-consuming method.
Key words: energy, industry 4.0, gas pipelines, energy transition, natural resources, non-destructive testing; life-cycle assessment, deep learning, convolutional neural networks, object detection.
DOI: 10.17580/cisisr.2024.01.15

Introduction

At present time gas pipelines are the main way for natural gas transportation, and their quantity only increases. Analysis of the works [1–4] shows that surface corrosion defects are met most often in gas pipelines during their operation; they are caused by the environment affect. These defects can be classified on plane ones (e.g. stress corrosion cracks) and volumetric defects, such as local pitting corrosion [5]. Several methods of non-destructive testing (NDT), which can reveal such defects automatically, are used in the industry. Visual methods as well as magnetic, eddy-current and ultrasonic methods are applied most often. Use of NDT methods separately can't provide collection of data set required for reliable classification of surface defects by their types, thereby combination of data received via different NDT methods is required to solve this problem. Joint processing of the results of ultrasonic and eddy-current control was examined in [6], use of Bayes conclusion and Dempster-Shafer theory are displayed in [7]; they are aimed on creation of a classifier for surface defects by their types. The observed NDT methods are characterized by false measuring results, while analysis of these results can be accompanied by false detection of large putting corrosion as the area without contact between transformer and gas pipeline metal. Appearance of pitting corrosion can be easily revealed visually, but in automatic mode a visual inspection is realized using TV control cameras. Analysis of pictures of the whole pipeline surface, which are obtained from TV control cameras, requires large volume of data set storage and essential labour efforts. In this connection, TV control system should be supplied by the model of computer vision, which can classify the images as snapshots without corrosion and snapshots with pitting corrosion, as well as to provide localization of pitting corrosion on the images.

At present time, there are a row of researches describing detection of external corrosion in pipelines using TV control equipped with computer vision function. This detection usually is accomplished with use of conventional algorithms of computer vision, machine and deep learning. These technologies were used for identification of corrosion defects on metallic surface [8–13].

Many kinds of architecture of convolutional neural networks (CNN) were suggested recently [14–19], they provide reliable classification of images based on the known data sets. Now deep CNN are used in various industries [20–24]; CNN is also applied for corrosion detection during processing of images [25–30]. This research is also devoted to development of the CNN-based model, which allows to identify corrosion of the pipeline surface; however the developed model is intended for classification of images without corrosion and images with pitting corrosion, as well as for its localization.

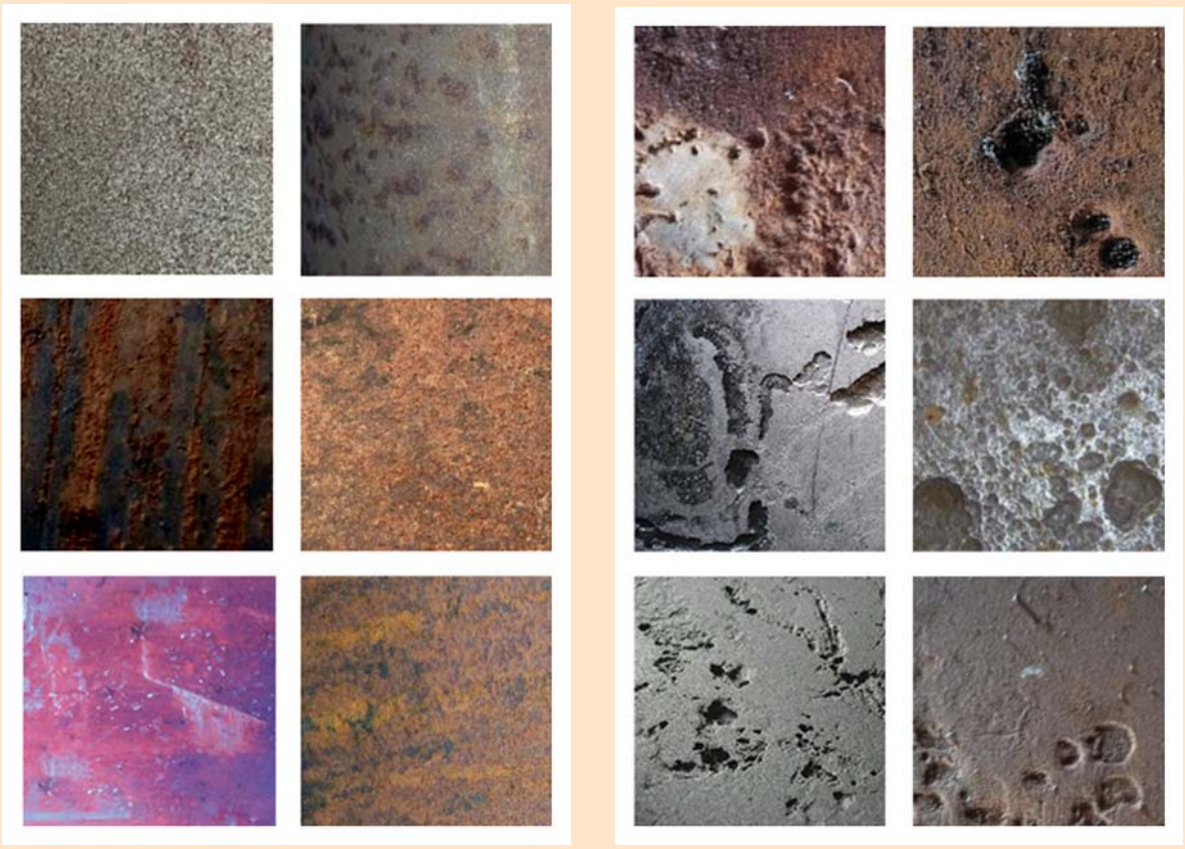
Rust and other variations of surface colour, when loss of wall thickness does not exceed 0.01 mm, are allowable for pipelines, because they don't decrease mechanical properties of a pipeline. Thus, rust with various colours and non-clean (not rusty) surface will be classified as images without corrosion. Images with pitting corrosion, with real danger of metal loss in a cavity, are classified as images with corrosion. This model was not found in the literature (most of models determine different rust degrees, as well as other defects), thereby the aim of this research is collection of the corresponding data, development of such model, its learning and efficiency assessment. Use of such system, allowing to identify and localize pitting corrosion on images together with ultrasonic and eddy-current NDT methods, as well as with use of laser triangular sensors, will provide reliable identification of type of the surface defects and to determine their parameters.

The technique of research and materials

Data set

The features of data set, as a rule, have maximal effect on the process of CNN learning. At present day, there is no assembled data set in open access, to provide classification of images with pitting corrosion and without corrosion, and obtaining of a rather large data set with pipeline surface images having pitting corrosion is not a simple task. Many pictures of defects on different metallic surfaces under various angles were made at many oil and gas industrial objects with non-destructive testing applied. All examined objects were manufactured from structural low-alloy steel or other low-carbon low-alloy steels with similar properties. The all these pictures were cut to the size 224 × 224 pix. The images from video shooting of TV surface control of pipelines, made from low-carbon low-alloy steels, were added to this data set afterwards. These images were also cut to the same size. At last, a row of pitting corrosion pictures were found in public data banks, such as www.shutterstock.com and www.dreamstime.com. Several images were got as a result of simple Internet search. The final data set consists of 5,760 images, where 4,270 pictures are images without corrosion and 1,490 images – with pitting corrosion. The examples of both classes are shown in the **Fig. 1**.

During the process of model development, the data set was divided to learning, validating and testing sets with random change of images, but the relative amount of each image class in total data set was the same. However, testing and validating sets have the same size and occupy totally 20 % of the whole data set. Additionally, the examined data set was characterized by imbalance of classes; there were images without corrosion approximately by 2.87 times more than images with pitting corrosion.



(*a*) (*b*)

Fig. 1. Examples from the data set of images without corrosion (*a*) and with pitting corrosion (*b*) on the surface of pipelines made from low-carbon and low-alloy steel

Model synthesis

At present time, use of deep learning technologies is one of the most efficient practice for image analysis. In particular, CNN as deep learning models present dominating approach for classification of images, reveal of objects and other tasks of computer vision. There are many kinds of CNN architecture today, such as AlexNet [14], ZFNet [15], VGGNet [16], Inception [17], ResNet [18] and Xception [19]. Use of the models which were preliminarily learnt by ImageNet data set is not recommended, because examined images are seriously varied from the images of ImageNet data set.

The models based on the above-mentioned architecture were learnt from scratch in the examined data set, and the best results were observed for the model based on ZFNet architecture. To provide maximal reliability during classification, self architecture based on ZFNet base was built. Original ZFNet architecture has its own flatten layer and two fully connected layers with 4096 parameters. However, fully connected layers are tended to overfitting, thereby the developed CNN has a layer of global average pooling (GAP) [31], which decreases number of parameters and possibility of the overfitting. The number of filters is also diminished in the developed CNN: 64 in the first two convolutional layers and 128 in other convolutional layers. Different modifications were tested manually, but this modification displayed the best results. To find the most optimal architecture of the developed CNN, automatic search of its hyper-parameters was carried out. Use of such widely developed methods as grid and random search was not suitable in this case, because learning of the examined architecture takes a lot of time and calculating resources, while larger number of iterations is required for these searching methods. In this case, Bayes optimization was used [32], applying different surrogate functions, such as Gauss process (GP) [33], randomized forest (subsequent model-based algorithm configuration – SMAC) [34] and tree-structurized Parsen evaluation (TPE) [35].

When using Bayes optimization, the values of the aimed function were minimized relating to F1-score, because the data set has definite imbalance of classes. As a result, three developed architectures were obtained: initial self architecture, self architecture with GP optimization and self architecture with SMAC optimization. All developed architectures trained from scratch on the training component of the examined data set. Learning was conducted with use of RMSprop optimization algorithm, with learning rate equal to $1 \cdot 10^{-5}$. To prevent overfitting, L2 regulation with the parameter equal to 1×10^{4} was used, while dropout with possibility 0.5 was used in fully connected layers. As soon as data set includes images from different sources, all architectures are characterized by mini-batch normalization [36] before activation functions. A data set has imbalance of classes, that's why the function of losses was used with the following weights for classes; 1 for images without corrosion and 2.5 for images with corrosion. To provide the best separation of classes, the tool «Label smoothing» with 0.3 parameter was used. CNN learning was realized by GPU nVidia RTX 3090 with memory 24 GB, batch size was equal to 64. The whole required code was developed using Python 3.9 language, with assistance of generally available libraries TensorFlow (2.5.0) and OpenCV (4.5.1).

Various metrics were used to determine classifier reliability for a testing data set. Confusion matrix is one of the most widely used methods in machine learning. It displays information about true classes and classes predicted via a model. Confusion matrix has two dimensions: true and predicted classes. The lines present an example of true class, while the columns – calculated class. *TP* in the matrix of errors means number of true positive examples, *TN* – number of true negative examples, *FP* number of false positive examples and *FN* – number of false negative examples. Metrics for examination of classification model efficiency are presented by accuracy, F1-score, area under the curve of receiving set operating parameter and under the curve of accuracy and recall (*ROC AUC* and *P-R AUC* respectively).

Accuracy means a number of correctly classified examples, which is divided by total number of examples, it is calculated via the following formula (1):

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

F1 score is an average harmonic value of such parameters as precision and recall. Precision shows number of positive data which was predicted correctly. High precision means smaller number of false operations. Recall determines integrity of a classifier. Higher recall displays smaller number of false negative results, while smaller recall means larger number of such results. Precision decreases often with recall improvement [37]. Precision, recall and F1-score are calculated in the following ways (2) – (4):

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

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

$$F1score = \frac{2\ Recall\ Precision}{Recall + Precision} \quad (4)$$

The curve of receiving set operating parameter (*ROC*-curve) is a graph of relationship between frequency of true positive results (*TPR*) and frequency of false positive results (*FPR*) with different threshold values of a classifier. Area under the curve (*AUC*) is the main parameter of a *ROC*-curve. The higher is *AUC*, the better is account of possible threshold values by a classifier. *TPR* and *FPR* are calculated in the following way (5) – (6):

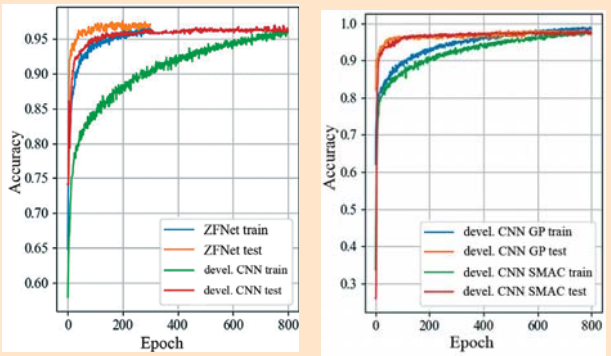
$$TPR = Recall = \frac{TP}{TP + FN} \quad (5)$$

$$FPR = \frac{FP}{FP + TN} \quad (6)$$

The curve of precision-recall (*P-R* curve) is used for efficiency evaluation of binary classification models in the same way as *ROC*-curve. It is often used in situations, when classes are not strongly balanced. The curve of precision-recall is built via preparing the graph of relationship between precision and recall for one classifier and for different threshold values. *AUC* for *P-R* curve, as *AUC* for *ROC*-curve,

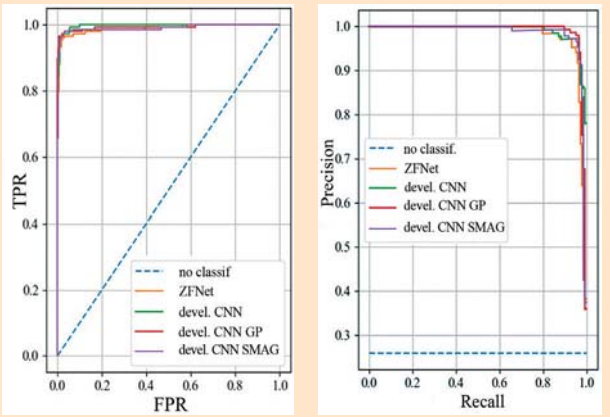
Table 1. **Reliability of developed architecture in comparison with ZFNet in a testing set**

Architecture	Accuracy, %	F1-score, %	*ROC AUC*	*P-R AUC*
ZFNet	96.70	95.30	0.98	0.98
Developed	97.40	96.97	0.99	0.99
Developed with GP optimization	98.44	97.30	0.99	0.99
Developed with SMAC optimization	97.40	96.30	0.99	0.98



(*a*) (*b*)

Fig. 2. Precision for developed architecture CNN and ZF Net for learning and testing data set: ZFNet and developed CNN (*a*); developed CNN with GP and SMAC optimizations (*b*)



(*a*) (*b*)

Fig. 3. *ROC* (*a*) and *P-R* (*b*) curves for developed architecture CNN and ZFNet

is a metrics within a range from 0 to 1. The higher is *AUC* for *P-R* curve, the better classifier operates.

Reliability of the developed models with optimization in a testing set in comparison with ZFNet architecture is shown in the **Table 1**. Accuracy graphs for each epoch, for the developed CNN architecture, are displayed in the **Fig. 2**. *ROC* and *P-R* curves for developed architecture can be seen in the **Fig. 3**.

The best accuracy and F1-score results were displayed by the developed CNN architecture with GP optimization, with values 98.44 % and 97.3 % respectively; its complete architecture is presented in the **Table 2**.

The developed CNN architecture has also smaller number of parameters in comparison with other well-known architectures. E.g., ZFNet architecture has about 30mln. parameters, VGG16 architecture – appr. 138 mln. parameters. The developed CNN architecture with GP optimization has only 2.4 mln. of learning parameters. Owing to smaller number of CNN parameters, class determination is quick and occupies less memory.

Experimental research and discussion

Reliability analysis of the developed model

Reliability analysis of the developed model was carried out for a testing data set, which was presented by images of pipelines from low-carbon and low-alloy steels. The calculation error matrix was built according to the developed CNN with GP optimization in the beginning of this analysis, it is shown in the **Table 3**.

Table 2. **Developed CNN architecture with GP optimization**

No.	Layer	Entrance	Filter size	Number of filters	Step	Exit	Number of parameters
1	Conv1	224 × 224 × 3	7 × 7 × 3	32	2	109 × 109 × 32	4736
2	Pool1	109 × 109 × 32	3 × 3 × 1	–	2	54 × 54 × 32	–
3	Conv2	54 × 54 × 32	7 × 7 × 32	128	2	24 × 24 × 128	200832
4	Pool2	24 × 24 × 128	3 × 3 × 1	–	2	11 × 11 × 128	–
5	Conv3	11 × 11 × 128	3 × 3 × 128	384	1	9 × 9 × 384	442752
6	Conv4	9 × 9 × 384	3 × 3 × 384	384	1	7 × 7 × 384	1327488
7	Conv5	7 × 7 × 384	3 × 3 × 384	128	1	5 × 5 × 128	442496
8	Pool3	5 × 5 × 128	3 × 3 × 1	–	2	2 × 2 × 128	–
9	GAP	2 × 2 × 128	–	–	–	128	–
10	FC	128	–	1	–	1	129

Table 3. **The calculation error matrix according to the developed CNN with GP optimization for a testing data set**

True class	Without corrosion	424	3
	Pitting corrosion	6	143
		Without corrosion	Pitting corrosion
		Class determined by a model	

It can be seen that the model correctly determined the class of majority of images (567 pictures). However, there were also mistakes: 6 images with pitting corrosion were classified as the images without corrosion, while 3 images without corrosion were identified as images with pitting corrosion.

Based on the results of conducted research, the CNN-based model was created; it allows to classify the images as images without corrosion and images with pitting corrosion on the surface of pipelines from low-carbon and low-alloy steels. In addition to identification of pitting corrosion in an image, it is usually required to provide localization of identified pitting corrosion. There are a lot of methods for detection of objects in images, which use conventional techniques of computer vision [38, 39] and deep learning [40–42]. However, these methods require a lot of data for learning, and these data are marked in accordance with a specialized format. This marking concludes not in classifying each image with a definite class, but in extraction of the area which corresponds to the examined object in an image and determination of its class. Such marking of a data set needs large labour expenses, as well as consequent learning with other more complicated models.

This research used the method of "sliding windows", when the preset area, presented by a triangle with preset number of pixels, moves along the image with definite step. The image obtained in this "window" is classified during each step using the developed CNN-based classifier. The method of "sliding windows" was applied together with the "pyramid of images", and image scaling occurred as a result in addition to "window" moving. As a result, a row of objects identified as pitting corrosion, were detected in the examined image. Integration of the detected areas occurs using non-maximum suppression technology (NMS) [43]. The general scheme of localization algorithm and identified pitting corrosion is presented in the **Fig. 4**.

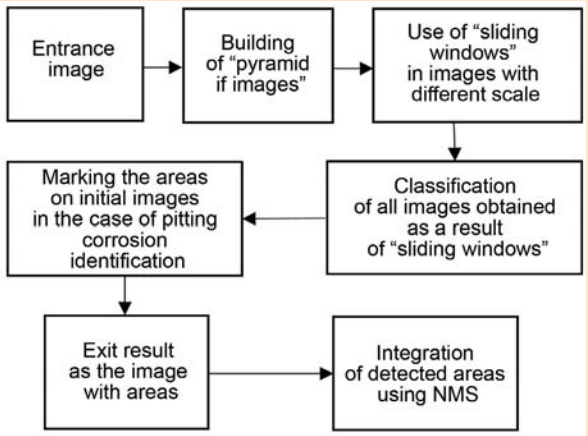


Fig. 4. General scheme of localization algorithm for identified pitting corrosion

The examples of localization of identified pitting corrosion before and after NMS are displayed in the **Fig. 5**.

It can be seen from the **Fig. 5** that the system identifies and localizes pitting corrosion (areas with metal loss), as well as changes in texture of the main metal, welding seam and other contaminations, which have no effect on mechanical properties of pipelines from low-carbon and low-alloy steels. Calculation of number of identified objects and their summarized square allows to evaluate number of pitting and caverns per surface unit, what is necessary for carrying out calculations of surface.

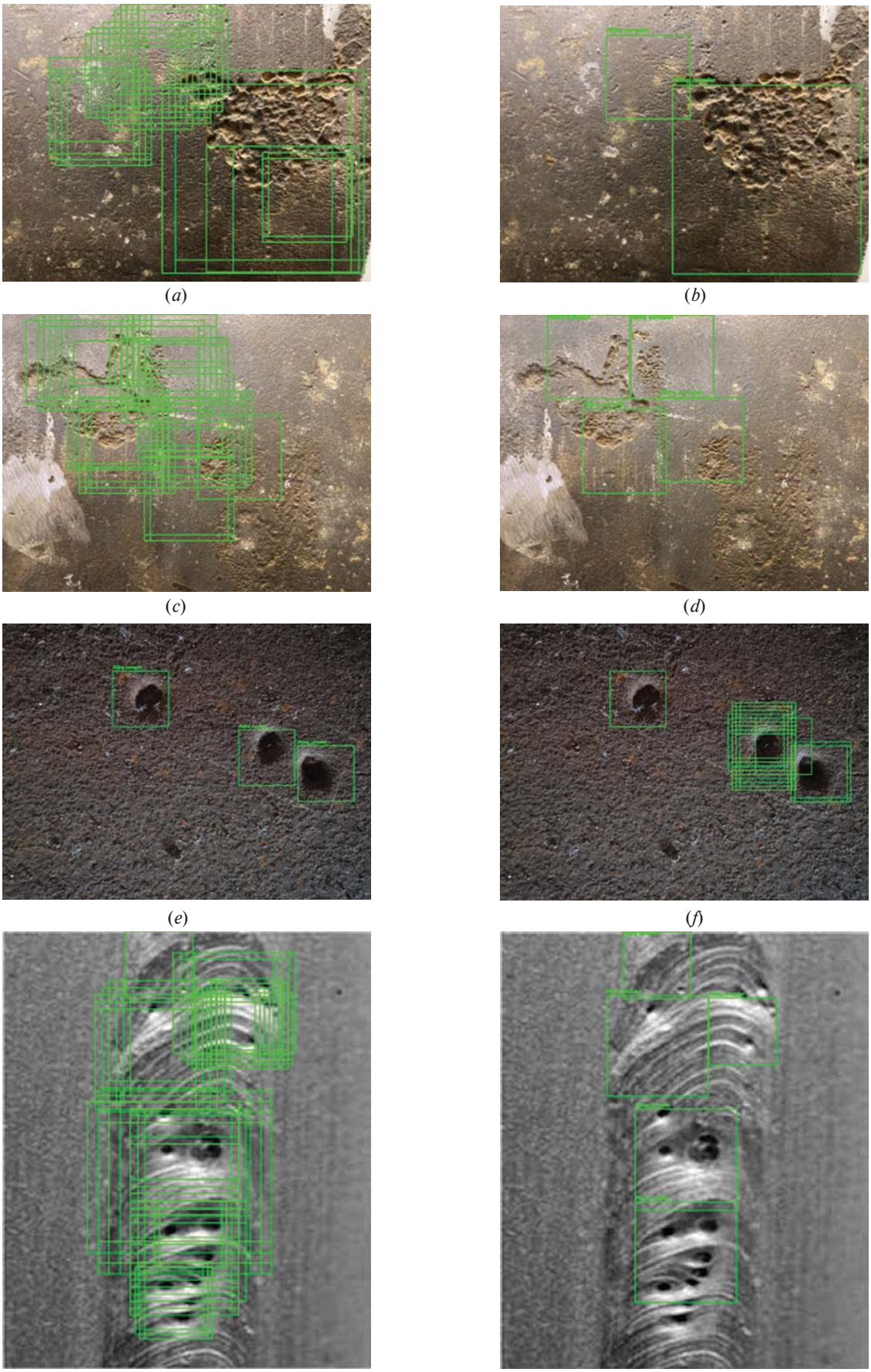
Conclusion

Surface defects, such as pitting corrosion, which can be easily detected via visual control, are forming on the external surface of gas pipelines manufactured from low-carbon and low-alloy steels. TV control cameras are usually used for automatic visual control. Manual saving and analysis of images from the whole pipeline surface is very labour-intensive process. In this connection, TV control systems are usually equipped with computer vision models, which can conduct identification and localization of pitting corrosion on metallic surface. The set including 5,760 images without corrosion and with pitting corrosion was created in this research. Self architecture on the base of well-known ZFNet architecture was also developed; consequently it was optimized using Bayes optimization with surrogate functions of GP and SMAC types. GP-optimized architecture displays the best results with accuracy 98.4 %. The developed architecture has less number of parameters and operated rather better than other CNN architectures. The scheme for localization of identified pitting corrosion in images was created. These images were received from the surface of pipelines manufactured from low-carbon and low-alloy steels, using the method of "sliding windows" and "pyramid of images".

The research was carried out with financial support of the grant of Russian Scientific Fund No. 22-29-00524, https://rscf.ru/project/22-29-00524/.

REFERENCES

1. Davydova D. G. Defects of technological pipelines: typology, evaluation of the effect on operation. *Prombezopasnost-Priuralye*. 2012. No. 8. pp. 24-28.
2. Erekhinskiy B. A., Maslakov S. V., Shustov N. I., Mitrofanov A. V., Baryshev S. N., Zaryaev M. Yu., Kravtsov A. V., Egorov S. V. Metal cracking in Christmas-tree gate valves of gas producing wells in northern deposits. *Territoriya Neftegaz*. 2014. No. 2. pp. 31–36.
3. Safina I. S., Kauzova P. A., Gushchin D. A. Evaluation of technical state of vertical steel tanks. *TekhNadzor*. 2016. No. 3 (112). pp. 39–42.



(*a*) (*b*) (*c*) (*d*) (*e*) (*f*) (*g*) (*h*)

Fig. 5. Examples of localization of identified pitting corrosion: before NMS (*a, c, e, g*) and after NMS (*b, d, f, h*)

4. Butusov D. S., Egorov S. I., Zavyalov A. P., Lyapichev D. M. Corrosion cracking of gas pipelines under stress: A manual. Moscow: Izdatelskiy tsentr RGU nefti i gaza imeni I. M. Gubkina. 2015. 80 p.
5. Kalinichenko N. P., Vasilyeva M. A. Atlas of defects in welding joints and main metal: A manual. Tomsk: Izdatelstvo Tomskogo politekhnicheskogo universiteta. 2006. p. 55.
6. Aleshin N. P., Skrynnikov S. V., Krysko N. V., Shchipakov N. A., Kusyi A. G. Classification of surface defects in the main metal of pipelines based on the results of complex diagnostics. *Kompyuternaya optika*. 2023. Vol. 47. No. 1. pp. 170–178. DOI: 10.18287/2412-6179-CO-1185.
7. Challa S., Koks D. Bayesian and Dempster-Shafer fusion. *Sadhana*. 2004. Vol. 29. pp. 145–176. DOI: 10.1007/BF02703729.
8. Choi K.-Y, Kim S. Morphological analysis and classification of types of surface corrosion damage by digital image processing. *Corrosion Science*. 2005. Vol. 47 (1). pp. 1–15.
9. Medeiros F. N., Ramalho G. L., Bento M. P., Medeiros L. C. On the evaluation of texture and color features for nondestructive corrosion detection. *EURASIP J. Appl. Signal Process*. 2010. No. 1. 817473.
10. Khayatazad M., De Pue L., De Waele W. Detection of corrosion on steel structures using automated image processing. *Developments in the Built Emvironment*. 2020. No. 3. 100022.
11. Ortiz A., Bonnin-Pascual F., Garcia-Fidalgo E. et al. Vision-based corrosion detection assisted by a micro-aerial vehicle in a vessel inspection application. *Sensors*. 2016. Vol. 16 (12). 2118.
12. Ahuja S. K. Surface Corrosion Detection and Classification for Steel Alloy using Image Processing and Machine Learning. *Helix*. 2018. Vol. 8 (5). pp. 3822–3827.
13. Hoang N. D., Tran V. D. Image processing –based detection of pipe corrosion using texture analysis and metaheuristic-optimized machine learning approach. *Hindwi. Computational Intellegence and Neuroscience*. 2019. article ID 8097213. 13 p.
14. Krizhevsky A., Sutskever I., Hinton GE. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*. 2012. p. 1097–105.
15. Zeiler M. D., Fergus R. Visualizing and understanding convolutional networks. *European conference on computer vision*. Springer. 2014. pp. 818–833.
16. Simonyan K., Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *The 3rd International Conference on Learning Representations (ICLR 2015)*. 2015. pp. 1–14.
17. Szegedy C., Vanhoucke V., Ioffe S., Shlens J., Wojna Z. Rethinking the Inception Architecture for Computer Vision. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA*. 2016. pp. 2818–2826.
18. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA*. 2016. pp. 770–778.
19. Chollet F. Xception: Deep Learning with Depthwise Separable Convolutions. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA*. 2017. pp. 1800–1807.
20. Golyak I. S., Kareva E. R., Fufurin I. L., Anfimov D. R., Scherbakova A. V., Nebritova A. O., Demkin P. P., Morozov A. N. Numerical methods of spectral analysis of multicomponent gas mixtures and human exhaled breath. *Computer Optics*. 2022. Vol. 46 (4). pp. 650–658. DOI: 10.18287/2412-6179-CO-1058.
21. Wang L., Liu Y., Fu L., Wang Y., Tang N. Functional Intelligence-Based Scene Recognition Scheme for MAV Environment-Adaptive Navigation. *Drones*. 2022., No. 6. p. 120.
22. Fufurin I., Berezhanskiy P., Golyak I., Anfimov D., Kareva E., Scherbakova A., Demkin P., Nebritova O., Morozov. A. Deep Learning for Type 1 Diabetes Mellitus Diagnosis Using Infrared Quantum Cascade Laser Spectroscopy. *Materials*. 2022. No. 15. p. 2984.
23. Lobanova V., Slizov V., Anishchenko L. Contactless Fall Detection by Means of Multiple Bioradars and Transfer Learning. *Sensors*. 2022.,No. 22. p. 6285.
24. Bobkov A., Aung Kh. Real-Time Person Identification by Video Image Based on YOLOv2 and VGG 16 Networks. *Automation and Remote Control*. 2022. Vol. 83. No. 10. pp. 1567–1575.
25. Atha D. J., Jahanshahi M. R. Evaluation of deep learning approaches based on convolutional neural networks for corrosion detection. *Struct. Health Monit*. 2017. pp. 1110–1128.
26. Petricca L., Moss T., Figueroa G., Broen S. Corrosion detection using ai: a comparison of standard computer vision techniques and deep learning model. *The Sixth International Conference on Computer Science, Engineering and Information Technology*. 2016. pp. 91–99.
27. Ahuja S. K., Shukla M. K., Ravulakollu K. K. Surface corrosion grade classification using convolution neural network. *IJRTE*. 2019. Vol. 8. Iss. 3. September. pp. 7645–7649.
28. Bastiana B. T., Na J., Ranjithb S. K., Jijia C. V. Visual inspection and characterization of external corrosion in pipelines using deep neural network. *NDT and E International*. 2019. Vol. 107. 102134.
29. Katsamenis I., Protopapadakis E., Doulamis A., Doulamis N., Voulodimos A. Pixel-Level Corrosion Detection on Metal Constructions by Fusion of Deep Learning Semantic and Contour Segmentation. In: Bebis, G., et al. Advances in Visual Computing. *ISVC 2020. Lecture Notes in Computer Science*. Springer, Cham. 2020. Vol. 12509. pp. 160–169.
30. Zhang S., Li Z., Yang C., Zhu C. Segmenting localized corrosion from rust-removed metallic surface with deep learning algorithm. *Journal of Electronic Imaging*. 2019. Vol. 28. Iss. 4 (July). 043019.
31. Min Lin, Qiang Chen, Shuicheng Yan. Network in network. *ArXiv preprint arXiv*. 2013. 1312.4400.
32. Dewancker I., McCourt M., Clark S. Bayesian optimization for machine learning: A practical guidebook. *ArXiv*. 2016. 1612.04858.
33. Ebden M. Gaussian Processes for Regression: A Quick Introduction. *ArXiv*. 2015. Vol. 2. 1505.02965.
34. Hutter F., Hoos H. H., Leyton-Brown K. Sequential model-based optimization for general algorithm configuration. *International conference on learning and intelligent optimization*. 2011. pp. 507–523.
35. Bergstra J., Bardenet R., Bengio Y., Kégl B. Algorithms for hyper-parameter optimization. In: Shawe-Taylo J., Zemel R. S., Bartlett P., Pereira F., Weinberger K. Q. (Eds.). *25th Annual Conference on Neural Information Processing Systems, Granada (Spain), 12–15 December 2011*. 2011. pp. 2546–2554.
36. Ioffe S., Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *ICML'15: Proceedings of the 32nd International Conference on International Conference on Machine Learning*. 2015. July. Vol. 37. pp. 448–456.
37. Taner A., Öztekin Y. B., Duran H. Performance Analysis of Deep Learning CNN Models for Variety Classification in Hazelnut. *Sustainability*. 2021 Vol. 13. 6527.
38. Lowe D. G. Object recognition from local scale-invariant features. Computer vision, *The proceedings of the seventh IEEE international conference 1999*. 1999. pp. 1150–1157.
39. Felzenszwalb P., McAllester D., Ramanan D. A discriminatively trained, multiscale, deformable part model. Computer vision and pattern recognition. *IEEE international conference 2008*. 2008. pp. 1–8.
40. Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu C.-Y., Berg A. C. SSD: single shot multibox detector. *European conference on computer vision*. 2016. pp. 21–37.
41. Lin T.-Y., Dollár P., Girshick R., He K., Hariharan B., Belongie S. Feature pyramid networks for object detection. *IEEE Conference on Computer Vision and Pattern Recognition*. 2017. Vol. 1. pp. 2117–2125.
42. Ren S., He K., Girshick R., Sun J. Faster R-CNN: towards real-time object detection with region proposal networks. *Advances in neural information processing systems*. 2015. pp. 1137–1149.
43. Hosang J., Benenson R., Schiele B., Learning Non-maximum Suppression. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017*. 2017. pp. 6469–6477.