



Approach to weld segmentation and defect classification in radiographic images of pipe welds

V.A. Golodov^{a,*}, A.A. Maltseva^{b, **}

^a Almetyevsk State Oil Institute, 2 Lenin St., Almetyevsk, Russia

^b South Ural State University, 76, Lenin Prospekt, Chelyabinsk, Russia



ARTICLE INFO

Keywords:

Defect detection
Radiographic images
Neural networks
Pipe welds segmentation
Defect segmentation
Defect classification

ABSTRACT

Detecting indications (points suspicious for defects) in weld radiographs is an important research topic in the field of industrial non-destructive testing. Many computer-aided detection techniques have been designed for such applications as detecting indications occurrence, segmentation of the indication area, classification of the indications. However, these techniques are mainly focused on only one of the listed problems. Different defects may exhibit different visual properties in shapes, sizes, textures, contrasts, and positions, that often leading to ad-hoc solutions. The paper investigates to the fine tuning of the machine learning approach to high resolution radiograph processing produced by real-time radiography using digital detector arrays (DDA) method. The main contributions of this work are the preprocessing feature for human readability of the radiographic images and the proposed neural network-based solutions for all stages, from weld detection on the radiographic images and its segmentation to indication segmentation and classification. The designed approach was implemented as a web service with web site front-end. The demo version of the software is available with this instruction <https://github.com/NastyMitseva/DefectRecognitionSystem>.

1. Introduction

At pipe plants, applying computer vision for automated detection of weld defects is a matter of great demand [1]. Radiographic inspection is one of the most precise ways to detect defects in welds at early production stages. The quality of this inspection significantly affects the final percentage of defects. Missing defects at the early stages of production can cause considerable financial losses as many defects can be corrected in the first pilot batches of conveyor production.

The real-time radiography using digital detector arrays (DDA) technique inspection of pipe welds is a continuous process. The necessity for a human operator review and approval makes the television inspection of pipe welds a long-lasting process. The inspection process requires the operator's persistent concentration, which causes a substantial influence of the human factor on the operation [2].

As a solution to this problem, this paper proposes the automated approach to defect detection and defect classification of pipe longitudinal welded joints using radiographs. Among the potential research topics, this work investigates the welded joint detection on the

radiographs through DNNs (Deep Neural Networks).

A DNN is one of the deep-learning architectures. The term applies to an artificial neural network (ANN) with multiple layers between the input and output layers. The DNNs proposed for the image processing have many layers, hence the name "deep" networks. Most of used in the paper deep neural networks belongs to convolutional neural network (CNN, or ConvNet) that is a class of deep neural networks, most commonly applied to analyzing visual imagery. The CNNs are regularized versions of multilayer perceptrons.

Our primary motivation for using DNNs is their outstanding performance in image segmentation and classification tasks [3]. Besides, we also intend to research the usage of pre-trained DNNs in the context of welded joint segmentation and defect segmentation and classification.

We have 8 classes of defects under consideration. Here and below term defect applies to some indication on the image that may point to the weld to contain some lack. Of course, final decision have to be taken according to absolute sizes provided by document like the ISO 10 675 and internal factory standards. Defect detection and defect classification are processes of the radiograph processing when all indications are

* Corresponding author.

** Corresponding author.

E-mail addresses: golodova@agni-rt.ru (V.A. Golodov), anastasia.mittseva@gmail.com (A.A. Maltseva).

URL: <http://agni-rt.ru/en/> (V.A. Golodov), <https://www.susu.ru/en> (A.A. Maltseva).

segmented and classified.

The proposed algorithm of defect detection and classification is based on segmentation and classification neural networks. It consists of six subsequent stages:

1. Image preprocessing
2. Segmentation of the weld area
3. Weld classification
4. Defect segmentation for the weld with indications
5. Defect classification for the weld with indications,
6. Image post-processing

The algorithm showed high recognition accuracy sufficient for use in production for automatic detection and classification of defects of pipe welds.

The remaining of the paper is organized as follows. Section 2 comprises a brief presentation of the state-of-art algorithms for weld defect detection and classification by radiographic images. Section 3 describes the proposed method in detail. The experimental results obtained from different models and the comparison with related methods are presented in section 4. Section 5 describes software implementation issues. Finally, section 6 concludes the paper and suggests some directions for future research.

2. Related works

There are two groups of methods used to solve the problems of weld area segmentation and defect classification: classical computer vision algorithms and neural network methods.

2.1. Classic computer vision algorithms

V. Kalaiselvi et al. [4] describe the Computer-Aided Detection system, which is based on image processing techniques. The proposed approach consists of three stages: gradient image formation, filtration by Gaussian pyramidal filters algorithm, and segmentation by Expectation and Maximization algorithm.

Benzhi Chen et al. [5] apply the algorithm of defect detection with sparsity reconstruction of welds. Initially, the authors form a set of good welds. The formed set allows to extract features and to reconstruct background and weld. Subtraction of the reconstructed image from the original one demonstrates defects.

N. Nacereddine et al. [6] suggest a weld defect detection algorithm. Algorithm consists of 4 stages: preprocessing to improve image contrast, image segmentation, feature extraction, and defect classification. The defect classifier is based on a finite mixture model with the multivariate generalized Gaussian distribution (MGGD).

2.2. Deep neural network algorithms

Yuting Wang et al. [7] propose to use the pre-trained convolutional neural network RetinaNet for automatic weld defects detection and type identification. The training set consists of 6714 images with three defect types: blowhole, underfill or incomplete penetration, and tungsten inclusion. The proposed approach allows identifying the defect type and its position.

Bin Liu et al. [8] use the convolutional neural network VGG-16 for defect classification. The following three weld types are classified: blow holes or solid inclusions, transversal cracks, and none defects. The neural network has two scales of defect patches for input as the defect sizes are various.

Xinghui Dong et al. [9] apply the segmentation neural network U-Net to recognize defects in welds. Only patches containing defect areas are fed to the input in order to improve the sensitivity of the neural network. The authors also offer to replace the softmax layer with the decision tree classifier and use Maximally Stable Extremal Regions (MSER) to identify

candidate areas.

Wenhui Hou et al. [10] offer an algorithm for automatic defect recognition, the implementation of which consists of three stages. The first step is selecting a weld area using Otsu's method. At the second stage, a sparse auto-encoder network (SAE) is trained and tested on patches cut from radiographs. At the third stage, a sliding window algorithm detects defects in the entire seam area.

F. M. Suyama et al. [11] describe the algorithm aimed at detecting welded joints of oil pipelines on radiographic images obtained by means of the DWDI method. The proposed approach extracts information from the pipe area of the radiographic images to cut out patches and then uses a deep neural network VGG-VD-16 to determine which windows correspond to the welded joints.

N. Boaretto et al. [12] developed the method of automatic defect detection and classification based on radiographic images of welds obtained with the DWDI method. At the first stage, the area of the weld is isolated using median and average filters. At the next stage, the potential defects in the weld are detached, and the features of the found defects are selected using Wiener filter, top-hat operation, histogram equalization, and Otsu's method. The selected features are classified with a three-layer neural network.

Wenhui Hou et al. [13] also suggest using a neural network to classify defects in welds from radiographic images. The input of the neural network is fed with 32×32 patches cut out of the radiographic images. The neural network architecture consists of 4 convolution layers with activation function ReLU, 4 max-pooling layers, and 1 fully connected layer with 5 outputs for each defect type.

Li Yaping et al. [14] use a convolution neural network to classify suspected defective areas into 3 types: linear defect, round defect, and noise. The proposed neural network consists of 4 convolutional layers with ReLU activation function, 4 max-pooling layers, and 2 fully connected layers.

N. Yang et al. [15] describe an algorithm for classifying welds with defects based on a modified LeNet-5 neural network named CNN-Xray. The CNN-Xray architecture consists of 4 convolutional layers with activation function LReLU + Softplus, 2 max-pooling layers, and 1 fully connected layer. The CNN-Xray input is fed with 60×60 windows cut out of the radiographic images. As a result of testing, the authors have proven that CNN-Xray is superior to LeNet-5, ANN, and SVM methods in terms of recognition accuracy.

F. C. Chen et al. [16] describe the NB-CNN algorithm for the detection of weld cracks. A convolution neural network, in combination with the Naive Bayes classifier, is used for the detection of defects to reject false positives. The neural network consists of 4 convolutional layers with ELU activation function, 4 max-pooling layers, and 2 fully connected layers.

Since most existing detection technologies cannot simultaneously detect and classify a lot of defect types with high accuracy, we developed a new approach for defect detection and classification.

The background analysis of existing solutions shows that usage of neural network technologies leads to higher recognition accuracy, so in our work, we use deep convolutional neural networks for segmentation and classification of welds and indications.

3. The proposed approach

3.1. Novelty and industrial orientation

Since neural networks outperformed human in a lot of specific visual data processing tasks its industrial application became matter of time. But even in 2021 we have artificial intelligence only in weak stage, mean that only in some specific conditions artificial intelligence could overcome human and all existed approaches can not provide better results in noisy weakly predictable cases.

That is why robotic lines released in areas like automated robotic assembly line. Robotic line decreases variability of the process and in

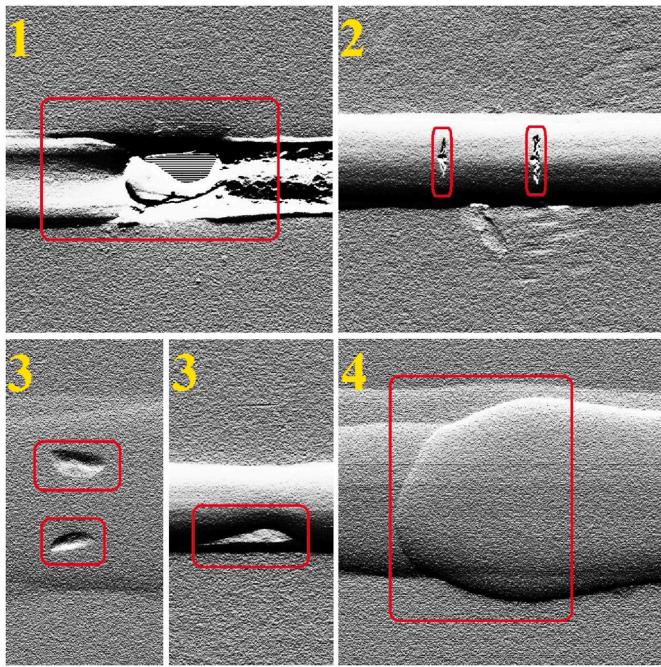


Fig. 1. Examples of the defect of types 1, 2 3, 4.

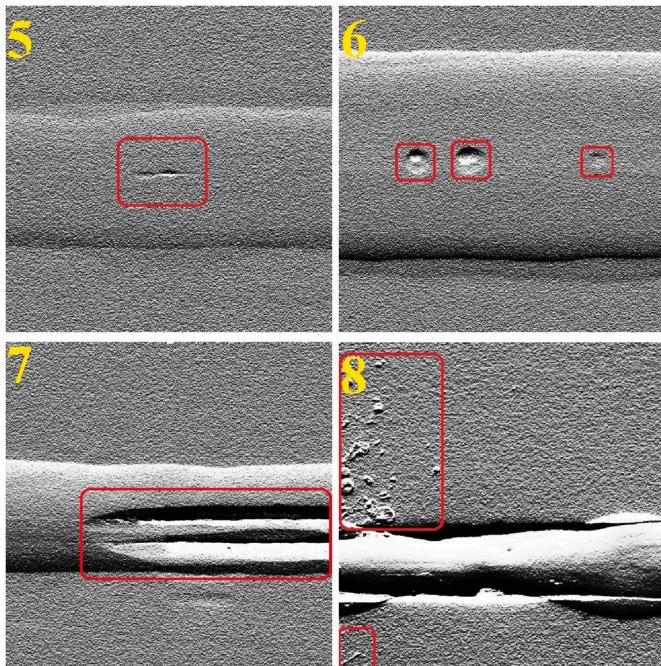


Fig. 2. Examples of the defect of types 5, 6 7, 8.

such decreased area of conditions become best choice. Another examples of lower local efficiency and higher global efficiency are multi-functional machines capable for entire complex process like CNC machines. Single subroutines may be performed in more efficient way but calibration time and calibration quality wastes all benefits.

For the task of the detecting defects in the pipe welds we have similar situation. At one side welding of the pipes is usually automated process performed with appropriate machine that decrease defect variability and on the other side we have to provide an approach to programming tool construction where all stages, may be not necessary in some cases, provides overall quality of the final tool.

So our approach consider problem of the software development for the defect detections on the pipe welds as tool construction process where we also consider our findings for best implementation of each part.

3.2. The proposed approach essence

The research aims to propose an approach to the problem of automatic classification of the defects of pipe welds with 8 classes according to ISO 6520:

1. Burn-through
2. Crack
3. Lack of fusion or undercut
4. Overlap
5. Incomplete penetration
6. Slag inclusion and pores
7. Weld breaking (not ISO 6520 naming)
8. Spatter.

Figures Fig. 1 and Fig. 2 show an examples of the pipe welds samples numbered with defect type number in the list above.

In this paper, a promising approach is proposed for weld defect detection and classification, which consists of six stages: image preprocessing, weld segmentation, weld classification, defect segmentation, defect classification, and image postprocessing. The following sections describe those stages.

3.2.1. Image preprocessing

The real-time radiography using digital detector arrays (DDA) is one of the most exact and high-technology way to find pipe weld defects. This work based on images provided by the defectoscopy unit completed with digital sensor marked as XRD 0822 AP3 IND. Sensor gives an image with resolution 1000×1000 pixels where one pixel corresponds to $200 \mu\text{m}$ area. After software adaption raw data are scaled to 1152×1152 16-bit grayscale image. The 1152×1152 images are source data used in the research.

Defectoscopy unit software encrypts all data to a ‘.VRC’ container intended for proprietary software decoding. In order to allow common purpose image processing libraries to edit encrypted data following conversion algorithm to ‘.png’ image series was developed as the first stage of the defect classification algorithm:

- file loads in binary mode;
- image raw data and additional position information are loaded;
- image raw data compose ‘.png’ image;

All the data are encoded 16-bit grayscale so require additional filter – intensity histogram normalization. The color intensity histogram is a chart where x-axis corresponds to intensity, and y-axis corresponds to number of pixels with given intensity.

Image normalization refines results of the deep learning methods. The waste data are avoided at the same time all significant data are scaled up. All radiographs given to authors have a strongly pronounced spike at one area and almost zero values for the other histogram area. Fig. 3 shows histograms of the same image before and after normalization procedure.

One of the most important stage of the successful image processing using neural networks is data collecting and preprocessing. High resolution data given by digital sensor allows to train machine learning algorithms such as DNN to find defects including small ones. Nevertheless human control and interpretation of the results are complicated for the high resolution gray-scale images including normalized ones.

The paper uses gradient processing and that reveals image data. The gradient is the vector formed by the operator ∇ (nabla) acting on a scalar function at a given point in a scalar field. The gradient image is directed

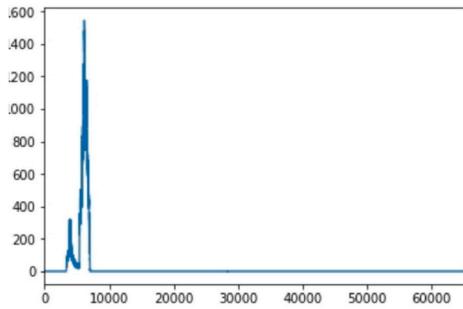


Fig. 3. Histogram of the original image (left) and its histogram after normalization by histogram (right).

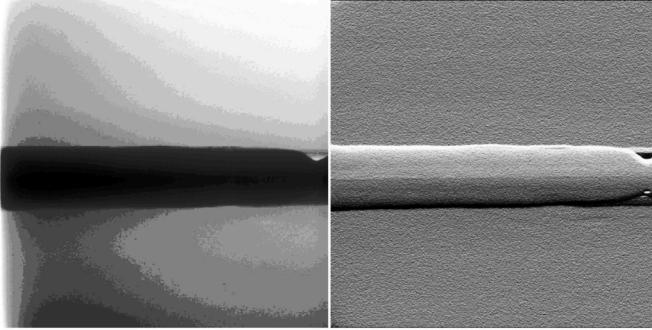


Fig. 4. The normalized image (left), the image normalized and preprocessed by gradient (right).

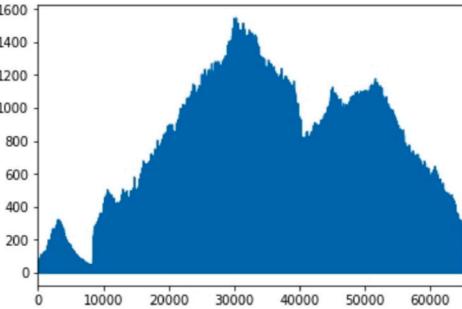


Fig. 5. Preprocessed image (left) and mask of its segmented weld (right).

Table 1
Weld segmentation models. Symbol “*” replaces the word FgSegNet in the first column.

Model	Acc.	Prec.	Recall	F1	ROC	IoU
*_S	0.966	0.921	0.913	0.917	0.946	0.847
*_M	0.980	0.967	0.937	0.952	0.964	0.908
*_v2	0.987	0.940	0.954	0.969	0.975	0.940
Unet	0.971	0.923	0.911	0.928	0.968	0.890
Linknet	0.966	0.915	0.940	0.918	0.967	0.883
PSPNet	0.961	0.945	0.865	0.890	0.963	0.852
FPN	0.947	0.911	0.878	0.878	0.943	0.833

color or intensity image transformation. Here the following discrete transformation (1) was used:

$$I_{g_{xy}} = \frac{-I_{x-2,y} + 8I_{x-1,y} - 8I_{x+1,y} + I_{x+2,y}}{12} \quad (1)$$

Fig. 4 shows normalized by histogram (left) and preprocessed using formula (1) image.

Right image on Fig. 4 allows both human and automated processing. All the data in this research are preprocessed in described way: histogram normalization plus gradient transformation. Further sections refer to the preprocessed image as an input image.

3.2.2. Weld segmentation

The second stage is the weld area segmentation to reduce noise and remove additional elements such as ruler and gauge. Experiments show that the neural network FgSegNet v2 [17] solves this task the best. The weld segmentation DNN selection experiments are described in section 4 see Table 1.

The FgSegNet_v2 is an neural network with an encoder-decoder structure. The first four VGG-16 [18] blocks are used as an encoder. Then the M_FPM block is connected, where feature maps of different scales and their concatenation are formed. The obtained latent space is fed to the decoder to create a binary mask. The neural network was

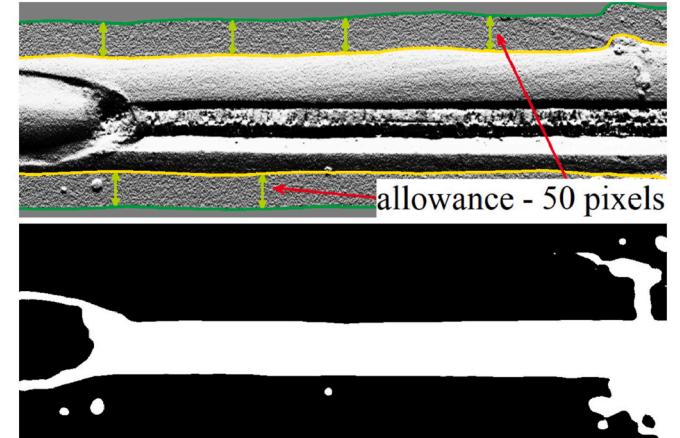


Fig. 6. Weld with indications adjusted after weld segmentation (upper) and mask of the segmented defects on the weld (bottom).

trained with RMSprop optimizer [19] with training speed 10^{-4} and batch size [20] equal to one.

After the weld area is segmented, the intermediate image processing is performed. Firstly, weld borders are defined, and we check if the neural network found the weld. It is required for possible cases of hard, very dark images where detection may fail. If a weld is detected successfully, a dimensional allowance of fifty pixels is added to capture the area near the weld, which also may contain defects.

Fig. 5 shows the result of the segmentation stage.

A weld with the adjacency is then cut from the image and scaled to 1152×384 , empty pixels are filled gray. The upper image on Fig. 6 shows the result after the extracting the weld with the adjacency.

3.2.3. Weld classification

The third stage is welds classification to smooth and welds with

Table 2

EfficientNet family results for the classification of the welds. Symbol '*' replaces the word EfficientNet in the first column.

Model	Acc.	Prec.	Recall	F1	ROC	Conf. matrix
*-B0	0.97	0.94	1.0	0.97	0.97	$\begin{pmatrix} 50 & 3 \\ 0 & 53 \end{pmatrix}$
*-B1	0.96	0.92	1.0	0.96	0.96	$\begin{pmatrix} 49 & 4 \\ 0 & 53 \end{pmatrix}$
*-B2	0.95	0.91	1.0	0.95	0.95	$\begin{pmatrix} 48 & 5 \\ 0 & 53 \end{pmatrix}$
*-B3	1.0	1.0	1.0	1.0	1.0	$\begin{pmatrix} 53 & 0 \\ 0 & 53 \end{pmatrix}$
*-B4	0.96	0.92	1.0	0.96	0.96	$\begin{pmatrix} 49 & 4 \\ 0 & 53 \end{pmatrix}$

Table 3

Defect segmentation models. Symbol '*' replaces the word FgSegNet in the first column.

Model	Acc.	Prec.	Recall	F1	ROC	IoU
*_S	0.987	0.901	0.876	0.888	0.935	0.799
*_M	0.990	0.907	0.916	0.911	0.955	0.838
*_v2	0.986	0.922	0.838	0.878	0.917	0.783
Unet	0.970	0.804	0.703	0.740	0.844	0.668
Linknet	0.977	0.834	0.714	0.768	0.893	0.654
PSPNet	0.972	0.794	0.602	0.667	0.945	0.585
FPN	0.958	0.884	0.654	0.720	0.857	0.599

Table 4

EfficientNet family results for the classification of the defects. Symbol '*' replaces the word EfficientNet in the first column.

Model	Top-1	Top-2	Prec.	Recall	F1
*-B0	0.827	0.967	0.71	0.72	0.71
*-B1	0.780	0.924	0.72	0.66	0.66
*-B2	0.791	0.956	0.76	0.75	0.75
*-B3	0.812	0.949	0.74	0.72	0.72

indications to reduce required computational resources and increase defect classification quality. Processing of the welds without indications finishes at this stage. The welds classification DNN selection experiments are described in section 4 see Table 2.

State-of-art neural network EfficientNet [21] is responsible for the weld classification. Here EfficientNet-B3 version was used with width coefficient 1.2, depth coefficients 1.4, and dropout equal to 0.3. Here width coefficient corresponds to the number of filters in convolution layers, depth coefficient corresponds to number of the MBConv blocks, and dropout value is the probability of neuron deactivation during training. EfficientNet-B3 is completed by the global average pooling layer [22] with a dropout equal to 0.2 and fully connected layer with one output and sigmoid activation function.

The RMSprop optimizer was used to train EfficientNet-B3 with the learning rate 2×10^{-5} and batch size equal to 2 during 50 epochs [20].

3.2.4. Defect segmentation

The fourth stage is the segmentation of the defects, and FgSegNet_M [23] shows the best results. Additionally, our neural network has four leading blocks of VGG-16 [18] pre-trained on ImageNet [24] as the encoder, where the fourth block is unfrozen during further training. The defect segmentation DNN selection experiments are described in section 4 see Table 3.

The FgSegNet_M gets the image in three scales as input, then latent codes from the VGG-16 are calculated. The received feature maps are adjusted to one scale, and this latent space follows into the decoder where the mask of the defects is formed.

The RMSprop optimizer was used to train EfficientNet-B3 with the learning rate $2e - 4$ and batch size equal to 1 during 50 epochs. Fig. 6

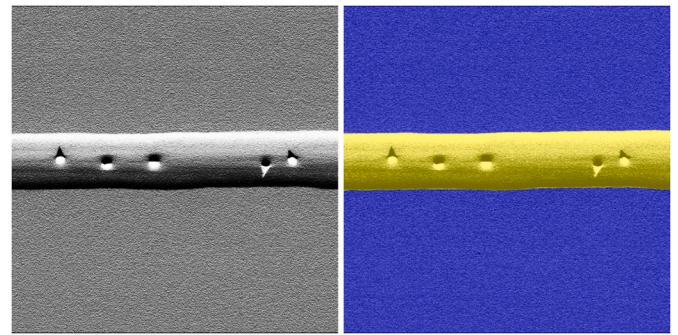


Fig. 7. Input image (left) and render of the weld segmentation, yellow for the weld. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

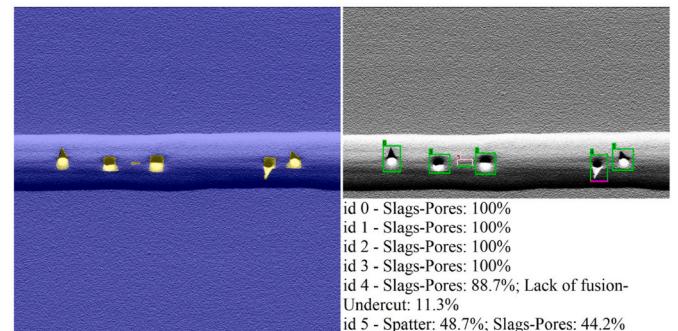


Fig. 8. Defect segmentation and defect classification results.

shows the result of the defect segmentation stage.

3.2.5. Defect classification

The fifth stage classifies the segmented defects. See Table 4 of section 4 about the defect classification DNN selection.

All defects are enclosed with the bounding boxes by scikit-image library functions and adjusted to 256×256 size. Next, the defect to one of the following 8 classes: 1) Burn-through, 2) Crack, 3) Lack of fusion or undercut, 4) Overlap, 5) Incomplete penetration, 6) Slag inclusion and pores, 7) Weld breaking (not ISO 6520 naming), 8) Spatter.

Efficientnet-B0 classifier is used with a width coefficient 1.0, depth coefficients 1.0, and a dropout equal to 0.2. It is completed with a global average pooling [22] with a dropout equal to 0.1 and a fully connected layer with 8 outputs and softmax activation function [25].

The RMSprop optimizer was used to train EfficientNet-B0 with the learning rate 2×10^{-5} and batch size equal to 8 during 50 epochs. During the training augmentation [26] and weight balancing [28] to deal with imbalanced classes [46] was used.

3.2.6. Image postprocessing

The last sixth stage is image postprocessing and forming on the final results. All the masks and the bounding boxes are adjusted to the original 1152×1152 image. There are several result renders to each input image for user assistance. Yellow color means area of concern violet color means exterior area.

First render corresponds to mask of the weld put on input image. Fig. 7 shows input one and the render of the segmented weld.

Second render contains mask of the defects found on the weld. This is useful for additional human control of the result if necessary.

Third render contains bounding boxes of the defects. Each bounding box has a unique id and is colored with two colors: main top-1 class color and additional top-2 class color. Only one color is used for 100% top-1 class cases. Image bottom contains the id of the defects and their classes.

		Prediction outcome		
		positive	negative	
Actual value	positive	TP	FN	TP + FN
	negative	FP	TN	FP + TN
		TP + FP	FN + TN	

Fig. 9. Confusion matrix.

Figs. 7 and 8 show the result of the sixth stage.

4. Experiments

This part of the paper presents experiments with different segmentation and classification models. The following tables contain a comparison with the state-of-art methods to demonstrate the effectiveness of the proposed algorithm.

Training and testing of all neural networks were passed on Windows@server with the NVIDIA@GeForce GTX 1080 8 GB and Intel(R) Core(TM) i7-6850K CPU.

4.1. Evaluation metrics

Common evaluation metrics for the segmentation and classification models are accuracy, recall, precision, F1-score, ROC AuC, IoU, and confusion matrix.

These metrics allow us to choose the best models and to compare different approaches to each others.

The *accuracy* metric is a basic metric that is a proportion of true results among the total number of cases examined.

The essential lack is inapplicability to imbalances classes cases. The accuracy metric [42] is calculated by formula 2.

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

where

- *TP*(TruePositive rate) – observation is positive and is predicted to be positive,
- *FN*(FalseNegative rate) – observation is positive but is predicted to be negative,
- *TN*(TrueNegative rate) – observation is negative and is predicted to be negative,
- *FP*(FalsePositive rate) – observation is negative and is predicted to be positive.

The *confusion matrix* [30] that demonstrates these factors is shown in Fig. 9. Same principles are used in multiclass case.

The *precision* [31] metric shows what proportion of the predicted Positives is truly Positive.

The *recall* [31] metric shows what proportion of actual Positives is correctly classified.

Formulas 3 and 4 demonstrate *precision* and *recall*, respectively:

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

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

A clear assessment of the machine learning algorithm quality gives ROC curve [31].

The *receiver operating characteristic curve (ROC)* [31] is created by plotting the true positive rate (TP) on the y-axis against the false positive rate (FP) on the x-axis at various threshold settings. Area Under Curve (AUC) is a numerical quality indicator. Perfect algorithm has ROC AUC equal to 1, while a bad one tends to 0.5.

F1-score is the harmonic mean of *precision* and *recall* and is used for the overall evaluation of the classifier quality. *F1-score* is calculated by the formula

$$F1 = 2 \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

Intersection over Union [32] (IoU) metrics is a ratio where the numerator is the area of overlap between the predicted bounding box and the ground-truth bounding box. The denominator is the area of the union, or more simply, the area encompassed by both the predicted bounding box and the ground-truth bounding box. IoU is calculated by the formula

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}}. \quad (6)$$

F1-score is a metric for the comparison between different papers and all above metrics are used during the choice of the best models for each stage.

4.2. Dataset

The data consists of the radiographs of large diameter pipe welds from the Chelyabinsk pipe factory. Manual image labeling was performed with COCO Annotator [47] free software. All labels were approved by the experts from the manufacture, who are radiographic unit operators.

The assigned machine learning tasks require different datasets. The dataset for the weld segmentation and classification contains 1060 images 1152×1152 pixels, number of smooth welds is equal to number of the welds with indications. The dataset for the defect segmentation contains 530 images 384×1152 pixels. The dataset for the defect classification contains 1857 images adjusted to 256×256 pixels. The training process also uses data augmentation [26] procedure (horizontal/vertical flips and image scaling).

4.3. Selection of the best models

The most popular state-of-art image segmentation and classification techniques were tested during our approach invention. Testing uses 15% of the dataset not opened in the training sample. The training process includes tuning the model parameters, batch size, used optimizer, learning rate, epoch number, augmentation characteristics, etc. See below for testing results for various algorithm steps.

Recognition accuracy for the best models is presented in Table 1. Applicants for the best weld segmentation model award were FgSegNet_S, FgSegNet_M, FgSegNet_v2, Unet [34], Linknet [35], PSPNet [36], FPN [37]. All models tested on 160 images, FgSegNet_v2 shows best results with F1-score equal to 0.969.

The models from the EfficientNet family were tested for the classification of the welds to smooth and welds with indications. The test sample contains 106 images. EfficientNet-B3 shows the best results with F1-score equal to 1.0. The recognition accuracy of tested models is presented in Table 2.

The defect segmentation tests the same neural networks as weld segmentation. The best score belongs to the FgSegNet_M model with F1-score equal to 0.911. All results are presented in Table 3.

The models from the EfficientNet family were tested for the classification of the defects to 8 classes. The test sample contains 278 images of 8 classes. Light and fast EfficientNet-B0 shows the best results by top-1 and top-2 accuracy, with 0.827 and 0.967 correspondingly. More heavy network can not cope with small defects. Results are presented in

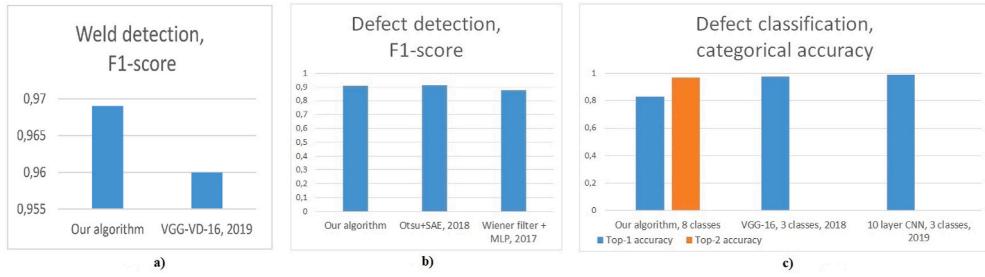


Fig. 10. Comparison with the state-of-art papers. (a) Different methods over the VARIOUS datasets and their F1-scores. In (a) VGG-VD-16: the method in Ref. [11] (b) Otsu + SAE:the method in Ref. [10]; Weiner filter + MLP: the method in Ref. [12] (c) VGG-16, 3 classes: the method in Ref. [8]; 10 layer CNN, 3 classes: the method in Ref. [14].

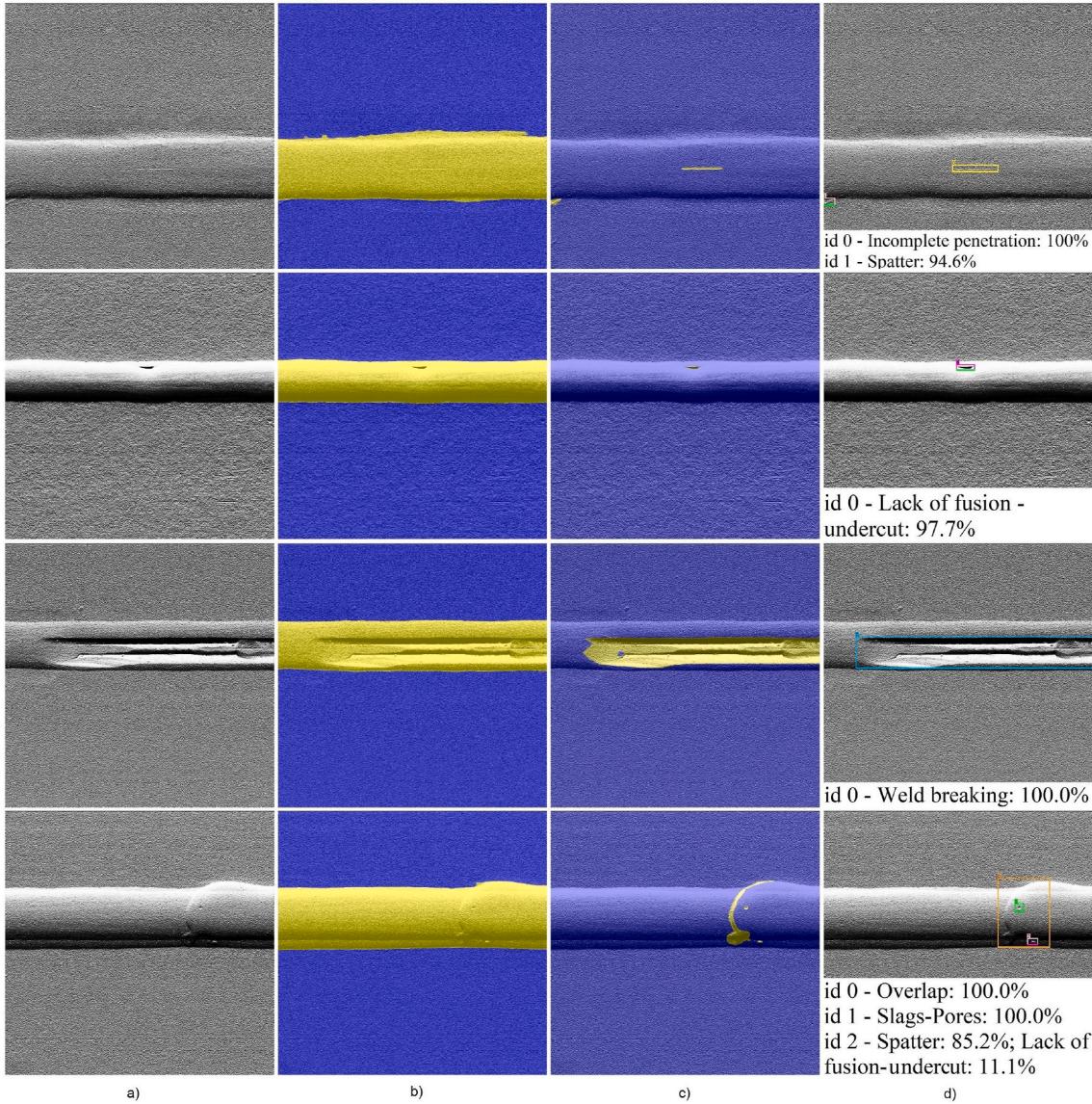


Fig. 11. Result renders samples 1–4. a) column – preprocessed input image, b) column – render of the weld mask, c) column – render of the defects masks, d) column – render of the defects classification results.

Table 4.

There are objective problems in the comparison of different approaches to weld segmentation, defect segmentation, and classification cited in the background section and newly developed one. All approaches use own training and testing sets of images from different

scanners. Therefore, the score provided below in Fig. 10 shows metrics of different approaches on their own test sets but allows us to compare the developed one with state-of-art ones published earlier.

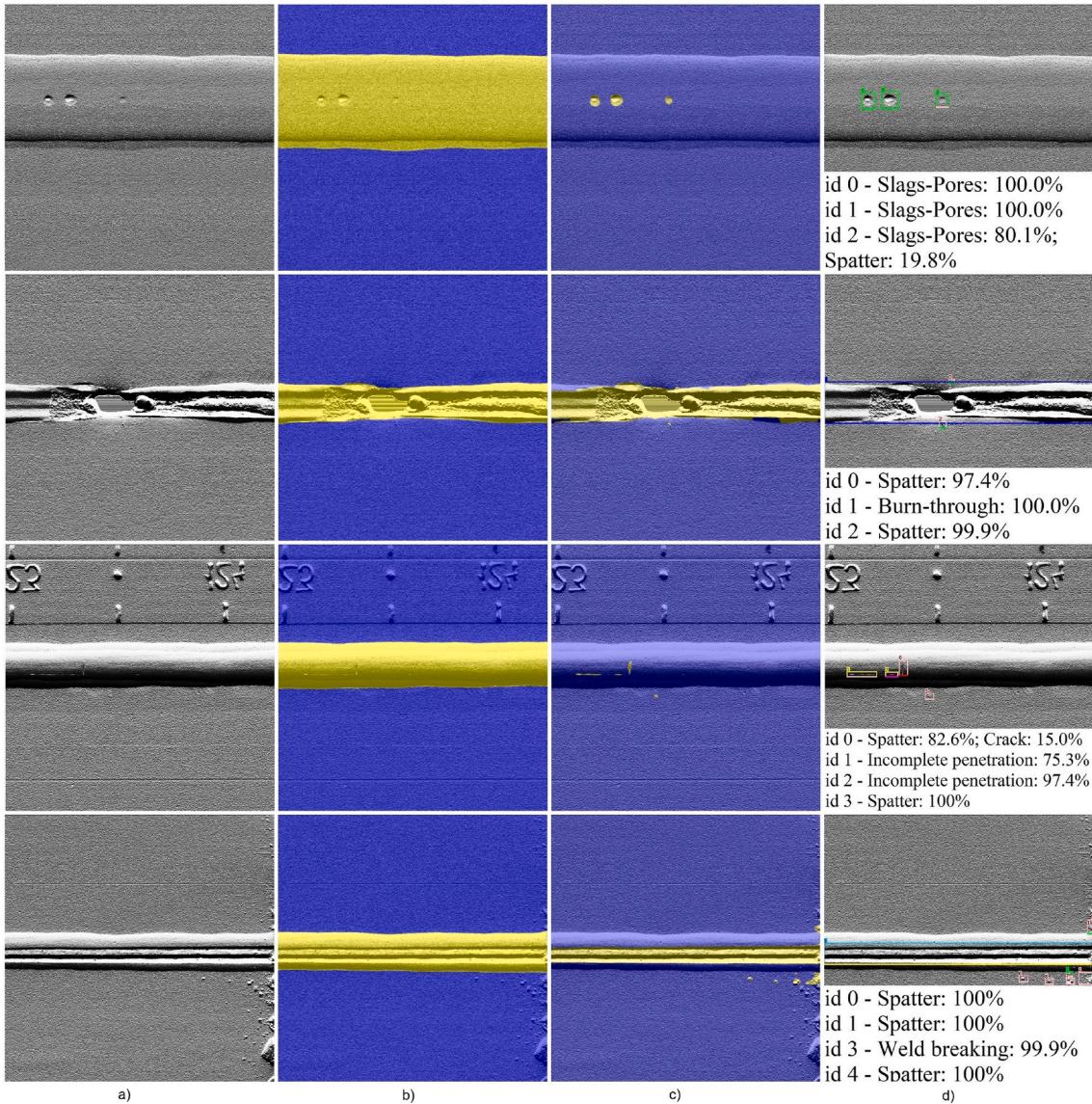


Fig. 12. Result renders samples 5–8. a) column – preprocessed input image, b) column – render of the weld mask, c) column – render of the defects masks, d) column – render of the defects classification results.

4.4. Discussion

This section contains results of the weld detection and classification, defect detection, and classification on the welds published in peer-reviewed journals. The introduced approach was compared with the existed ones according to F1-score and categorical accuracy metrics. Analysis shows [10] that the newly developed approach is competitive and exceeds accuracy of the lots of the existed ones.

The final results of the proposed approach on the testing data are:

- weld segmentation 96.94% F1-score,
- defect segmentation with 91.19% F1-score,
- defect classification: 82.73% top1 and 96.76% top2 categorical accuracy

The advantage of the newly presented approach is fast, automatic weld detection in an extensive range of image quality. The proposed defect segmentation approach handles small and huge defects in the same way. In the analyzed papers, there is a maximum of eight classes presented.

5. Software implementation

The proposed approach was implemented as a web service with web site front-end. Web service is flexible, providing REST API [38] for any client. Web site front-end provides user interface for a system operator. The user guide for the current software version is available at <https://github.com/NastyaMittseva/DefectRecognitionSystem>.

The web service was implemented using the Flask [43] micro-framework and MongoDB [44] database. All trained models have been deployed using Tensorflow Serving [45]. This service allows maintaining several models at the same time by addressing them via POST queries and simplifies the interaction between the web interface and the models.

Results from the developed software are presented in Fig. 11 and Fig. 12.

6. Conclusion

This paper has introduced a new approach for the defect segmentation and classification on the images from an real-time radiography

using a digital detector array (DDA). Our approach consider a software development process of the defects detection as tool construction. We also provide our analysis of the best candidates for all stages of the detection and classification process. The proposed approach consists of image preprocessing, weld detection, weld classification, defect detection, defect classification, and image postprocessing. The experiments show that weld segmentation with FgSegNet_v2 neural network has an F1-score 96.94%. Weld classification with Efficient-B3 neural network has 100% F1-score. Defect segmentation with FgSegNet_M neural network has an F1-score 91.19%. The final defect classification stage with Efficient-B0 gives 82.73% and 96.76% of categorical accuracy by top1 and top2 correspondingly. Further work is real-time image processing as part of the embedded algorithm for DDA unit.

CRediT authorship contribution statement

V.A. Golodov: Conceptualization of this study, Methodology, Resources, Validation, Supervision, Writing - Review Editing. **A.A. Maltseva:** Software, Data Curation, Formal analysis, Investigation, Writing - Original draft preparation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is supported by the Ministry of Science and Higher Education of the Russian Federation under the research grant no. FENU-2020-0022. We thank you for the radiographs provided by the Chelyabinsk pipe factory.

References

- [1] Hou W, Zhang D, Wei Y, Guo J, Zhang X. Review on computer aided weld defect detection from radiography images. *Appl Sci(Switzerland)* 2020;10:1–16.
- [2] Wall M. Human factors guidance to improve reliability of non-destructive testing in the offshore oil and gas industry. In: Proceedings of the 7th European-American Workshop on Reliability of NDE, Potsdam, 4-7 September 2017; 2017. p. 1–13.
- [3] Guo Y, Liu Y, Georgiou T, Lew MS. A review of semantic segmentation using deep neural networks. *Int. J. Multimed. Inform. Retr.* 2018;7:87–93.
- [4] Kalaiselvi V, Aravindhar DJ. An non destructive test for the detection of weld defects using image processing. *Indones. J. Electr. Eng. Comput. Sci.* 2018;9: 764–70.
- [5] Chen B, Fang Z, Xia Y, Zhang L, Huang Y, Wang L. Accurate defect detection via sparsity reconstruction for weld radiographs. *NDT E Int* 2018;94:62–9.
- [6] Nacereddine N, Goumeidane AB, Ziou D. Unsupervised weld defect classification in radiographic images using multivariate generalized Gaussian mixture model with exact computation of mean and shape parameters. *Comput Ind* 2019;108:132–49.
- [7] Wang Y, Shi F, Tong X. A welding defect identification approach in X-ray images based on deep convolutional neural networks, vol 10956. Springer International Publishing; 2019. p. 53–64.
- [8] Liu B, Zhang X, Gao Z, Chen L. Weld defect images classification with VGG16-based neural network, IFTC 2017. *Digital TV Wireless Multimedia Commun.* 2018;815: 215–23.
- [9] Dong X, Taylor CJ, Cootes TF. Small defect detection using convolutional neural network features and random forests. In: Computer Vision – ECCV 2018 Workshops, 11132; 2018. p. 398–412.
- [10] W. Hou, Y. Wei, J. Guo, Y. Jin, C. Zhu, Automatic detection of welding defects using deep neural network, in: *J Phys Conf*, pp. 1–10.
- [11] Suyama FM, Delgado MR, Dutra da Silva R, Centeno TM. Deep neural networks based approach for welded joint detection of oil pipelines in radiographic images with Double Wall Double Image exposure. *NDT E Int* 2019;105:46–55.
- [12] Boaretto N, Centeno TM. Automated detection of welding defects in pipelines from radiographic images DWI. *NDT E Int* 2017;86:7–13.
- [13] Hou W, Wei Y, Jin Y, Zhu C. Deep features based on a DCNN model for classifying imbalanced weld flaw types. *Measurement. J. Int. Meas. Confed.* 2019;131:482–9.
- [14] Yaping L, Weixin G. Research on X-ray welding image defect detection based on convolution neural network. *J Phys Conf* 2019;1237:1–7.
- [15] Yang N, Niu H, Chen L, Mi G. X-ray weld image classification using improved convolutional neural network. *AIP Conf Proc* 2018;1995:1–7.
- [16] Chen FC, Jahanshahi MR. NB-CNN: deep learning-based crack detection using convolutional neural network and Naïve Bayes data fusion. *IEEE Trans Ind Electron* 2018;65:4392–400.
- [17] Lim LA, Keles HY. Learning multi-scale features for foreground segmentation, pattern analysis and applications. 2018.
- [18] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-scale Image Recognition, CoRR abs/1409, 1; 2014. p. 1–14.
- [19] Ruder S. An overview of gradient descent optimization algorithms. 2016. p. 1–14.
- [20] Sharma Sagar. Epoch vs batch size vs iterations. 2017.
- [21] Tan M, Le QV. EfficientNet: rethinking model scaling for convolutional neural networks. In: 36th International Conference on Machine Learning, ICML 2019 2019-June; 2019. p. 10691–700.
- [22] Cook Alexis. Global average pooling layers for object localization. 2017.
- [23] Lim LA, Yalim Keles H. Foreground segmentation using convolutional neural networks for multiscale feature encoding. *Pattern Recogn Lett* 2018;112:256–62.
- [24] Takamitsu Y, Orita Y. Effect of glomerular change on the electrolyte reabsorption of the renal tubule in glomerulonephritis (author's transl). *Jpn J Nephrol* 1978;20: 1221–7.
- [25] Nwankpa CE, Ijomah W, Gachagan A, Marshall S. Activation functions: comparison of trends in practice and research for deep learning. *arXiv*. 2018. p. 1–20.
- [26] Buslaev A, Iglovikov VI, Khvedchenya E, Parinov A, Druzhinin M, Kalinin AA. Albumentations: fast and flexible image augmentations. *Information (Switzerland)* 2020;11.
- [28] King G, Zeng L. Logistic regression in rare events data. *J Stat Software* 2003;8: 137–63.
- [30] Ragan Allison. Taking the confusion out of confusion matrices. 2018.
- [31] Powers DMW. Evaluation : from precision , recall and F-factor to ROC , informedness , markedness & correlation. 2007.
- [32] Sheremet O. Intersection over union (IoU) calculation for evaluating an image segmentation model. 2020.
- [34] O. Ronneberger, P. Fischer, T. Brox, U-net: convolutional networks for biomedical image segmentation, in: Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics), volume vol. 9351, pp. 234–241..
- [35] Chaurasia A, Culurciello E. LinkNet: exploiting encoder representations for efficient semantic segmentation. In: 2017 IEEE visual communications and image processing, VCIP 2017 2018-Janua; 2018. p. 1–4.
- [36] Zhao H, Shi J, Qi X, Wang X, Jia J. Pyramid scene parsing network. In: Proceedings - 30th IEEE conference on computer vision and pattern recognition, CVPR 2017 2017-Janua; 2017. p. 6230–9.
- [37] Kirillov A, Girshick R, He K, Dollar P. Panoptic feature pyramid networks. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE; 2019. p. 6392–401.
- [38] Erl T, Carlyle B, Balasubramanian R. SOA with REST - principles, patterns and constraints for building enterprise solutions with REST. Prentice Hall; 2013.
- [42] Keras API reference / Metrics / Accuracy metrics. [Accessed 12 August 2021].
- [43] Flask's documentation. [Accessed 12 August 2021].
- [44] MongoDB. [Accessed 12 August 2021].
- [45] TensorFlow Serving. [Accessed 12 August 2021].
- [46] Estimate class weights for unbalanced datasets. [Accessed 12 August 2021].
- [47] COCO Annotator. [Accessed 12 August 2021].



Valentin Golodov, associate professor, senior researcher of the South Ural State University, senior researcher of the Almeteysk State Oil Institute. Valentin Golodov received the specialist degree (5-year education program) specialization in mathematical methods of forecasting in 2010 at Moscow State University. Candidate of Science since 2015, has wide scientific interests including but not limited to data mining, machine learning, computational mathematics. Current researches aim to neural networks application for different practical tasks, including industrial application.



Anastasia Maltseva (Mittseva) received the Master's degree in the Higher School of Electronics and Computer Science at the South Ural State University in 2020. From 2018 to 2019, she worked in the chelpipe group, and at the same time, she has started a development of the defect recognition system. She currently works at the company Dowell/Everypixel, and the current research is devoted to the development of generative networks for face swap and face reenactment tasks.