Journal of Innovations in Business and Industry

Vol. 04, No. 04 (2026) 425-436, DOI: 10.61552/JIBI.2026.04.006 - http://jibi.aspur.rs

THE DETERMINATION OF DEFECTS IN STEEL PIPES USING COMPUTER VISION METHODS BASED ON THERMOGRAPHIC IMAGES

Shcherban Pavel Sergeyevich [1]
Mazur Ekaterina Vladimirovna
Eranosian Sos Samvelovich
Fedor V. Piskun

Received 24.06.2026.
Revised 27.07.2026.
Accepted 28.08.2026.

Keywords:

Thermodiagnostics, steel pipes, pipe defects, defect segmentation, neural network modeling, online augmentation.

Original research

A B S T R A C T

Ensuring the reliability and safety of pipeline transportation systems in industry is a critical area of technical service. This is especially important when steam, hot water, oil, or other hazardous liquids are transported through pipes, as leakage can lead to serious emergencies. The development of new non-destructive testing methods and computer systems for analyzing data and detecting defects has significantly improved the accuracy and speed of pipeline inspections. For this research, a dataset of 210 thermographic images was collected from five types of cooling system pipes in a six-cylinder four-stroke engine. Existing methods for processing thermographic data were reviewed and neural network models were selected for training. The data was preprocessed and labeled. During dataset processing, computer vision methods based on convolutional neural networks were applied. The YOLOv8 neural network architecture was used. As a result, a neural network was developed that allows for the identification of volumetric and extended defects in thermographic images of pipes. In the determination of volumetric defects, an accuracy of 92.2% was achieved with a response time of 870 milliseconds. For extended defects, the mask accuracy was 58.6% and the response time was 800 milliseconds. The identification of point defects was hindered by interference in the thermograms, but a mask accuracy of 51.9% was still achieved with a response of 286 milliseconds.

1. INTRODUCTION

In recent decades, computer vision and machine learning techniques have been widely used in various scientific and technological fields, including industrial diagnostics. One of the main challenges in this area is automated processing of thermal imaging data, which is crucial for assessing the condition of industrial equipment and structures. Specifically, the analysis of thermal images of steel pipes is essential for timely detection of defects and prevention of accidents, as well as improving the operational reliability of machinery (Ma et al., 2021).

Traditional methods of analyzing thermal image data include the application of classical mathematical techniques such as image filtering, histogram transformation, and correlation analysis. These methods allow for preprocessing data, removing noise, and enhancing image contrast (Han et al., 2022). However, despite their effectiveness, these methods often require manual parameter adjustments and have limited flexibility when analyzing complex defects.

[1] Corresponding author: Shcherban Pavel Sergeyevich
Email: ursa-maior@yandex.ru

Modern machine learning and computer vision technologies offer new possibilities for the automatic analysis of thermographic data. Deep neural networks, particularly convolutional neural networks (CNNs), not only allow identifying defects in thermographic images but also classify them according to the type and severity of damage to equipment (Alzubaidi et al., 2021). The use of these algorithms significantly improves diagnostic accuracy and reduces human error, providing a high level of process automation (Ciampiconi et al., 2023; Liu et al., 2024).

One promising area is the development of software that can accurately process thermographic images of steel pipes while minimizing computational costs. It's important to consider the specific characteristics of these images, such as variations in temperature gradients, shooting conditions, and external noise. To solve these problems, we need to combine data preprocessing techniques with deep learning methods to ensure maximum accuracy and reliability in our analysis.

This research project aims to develop software for processing thermographic images of steel pipes using computer vision techniques (Golodov & Maltseva, 2022).. We will explore modern approaches to infrared image analysis, evaluate existing machine learning algorithms, and propose optimal solutions for automated defect detection.

Our main focus is on developing an effective method that ensures high accuracy in defect detection with minimal false positives. We aim to train a neural network and create a software tool that can automatically diagnose steel pipe defects by analyzing thermographic images using computer vision tools.

The practical value of this approach is related not only to the accuracy of defect detection, but also to its potential use in everyday industrial inspection. Automating the analysis of thermographic images can reduce the amount of manual work required from inspectors and make the inspection process faster and more consistent. Such a system could also provide useful information for maintenance planning and help identify pipe defects before they lead to more serious problems. In this way, the application of computer vision can support improvements in industrial quality control, maintenance, and the overall reliability of pipeline systems.

2. OVERVIEW OF VARIOUS METHODS OF THERMOGRAPHIC DATA PROCESSING AND JUSTIFICATION OF THE CHOSEN APPROACH

Classical mathematical methods occupy an important place in the processing of thermographic images, especially at the stages of data preprocessing, where noise reduction, contrast equalization and analysis of temperature gradients are required. These approaches rely on time-tested algorithms and mathematical models that ensure high accuracy and reliability of data processing. They are often used as a basis for the subsequent application of more complex methods or in cases where computing power resources are limited. Let's consider a number of classical and modern approaches to processing thermographic images.

Among the recent studies based on traditional methods, the work of Kim et al. (2022) stands out. In this study, the authors investigate the use of spatial filters for the preprocessing of thermographic images and the application of correlation techniques to quantify the extent of defects. The findings indicate that selecting an appropriate spatial filter for the pre-processing of thermographic data is crucial for subsequent analysis and assessment of defects. By combining a suitable filter with correlation analysis, high accuracy and dependability can be achieved in determining the size of flaws, which is essential for industrial diagnostics and nondestructive testing applications.

Another study (Liu et al., 2017) proposes a method for automatic crack detection in welded structures using pulsed eddy current thermography and image processing techniques (Alshawabkeh et al., 2025). The study focuses on the use of mathematical morphology-based algorithms and the automatic area growth method for analyzing thermographic data and combining images. Pulsed eddy current heating is used to heat the material, and defects, such as cracks, cause temperature anomalies in the heated area. These cracks appear as regions with different temperature gradients in the thermographic images.

At the same time, in recent years, there has been a steady increase in the range of methods for analyzing thermal imaging data due to the introduction of machine learning technologies. These modern intelligent algorithms are increasingly supplementing and, in some cases, replacing traditional approaches, offering more flexible, adaptive, and automated solutions. This is particularly important in the context of the increasing complexity of the objects being examined and the need for accurate diagnostics. Machine learning techniques allow not only minimizing the influence of the human element, but also identifying complex hidden patterns that might elude classical analysis methods. Thus, at the intersection of modern mathematical techniques and software tools, a new level of thermal image analysis is emerging, offering additional opportunities for nondestructive testing and industrial diagnosis.

These technologies make it possible to automate the process of analyzing thermographic data, significantly improving accuracy and reducing dependence on subjective factors. By using deep learning algorithms, such as convolutional neural networks (CNNs), it is possible not only to detect anomalies but also to classify them based on their context. In addition, machine learning integration ensures that algorithms are tailored to specific tasks, such as analyzing complex structural defects.

For example, the study by Qureshi et al. (2024) examines the use of radiometric infrared thermography (IRT) for diagnosing and predicting the maintenance of solar panels. The researchers propose an intelligent system for detecting and classifying defects based on lightweight convolutional neural networks (CNNs), which effectively solves the problem of multiclass classification.

The main focus of this publication is the development of a remote monitoring method to maintain stable energy production at solar power plants. To achieve high accuracy in diagnosis, the researchers created a special radiometric dataset with numerical floating-point temperature values that allows for taking into account thermal anomalies associated with solar panel degradation. The creation of this dataset was complicated by factors such as noise, the complexity of thermal patterns, and class imbalance.. The use of modern techniques of interpretable artificial intelligence helps to visualize the key characteristics of the data and understand the classification results. This enhances the reliability of the system and makes it more user-friendly for industrial applications.

Another study (Ullah et al., 2020) presents a novel technique for timely detection and diagnosis of defects in high-voltage equipment. Infrared thermography is used to effectively identify thermal anomalies caused by mechanical defects, such as cracks, damage to insulation, and contact problems, as well as overloads. This is crucial for preventing emergencies and maintaining stable operation of power systems.

Special attention is given to deep learning technology. At the data processing stage, the AlexNet model is used, known for its ability to extract spatial features. These features are then analyzed using machine learning algorithms, such as random forest and support vector machines. These algorithms demonstrate different levels of efficiency, with random forest showing the highest accuracy in classification, indicating its suitability for processing complex data.

Traditional image processing techniques include various filtering methods, segmentation algorithms, and morphological operations. For instance, the Gaussian filter is commonly used to remove noise and smooth images, although its use results in loss of detail and blurred borders. The median filter is more effective at handling pulse noise and preserving object boundaries, but as the filtering window size increases, so does the computational complexity. The averaging filter is simple to implement and reduces intensity variation, but it can also lead to blurred edges and lost fine details.

The correlation method demonstrates high accuracy in the presence of reference images, but requires their preliminary collection and storage, which increases the requirements for memory and computing resources. Morphological operations, such as erosion and dilation, make it possible to effectively isolate the structures of defects, but their result depends on correctly selected structural elements. Classical segmentation methods, including the Canny operator and threshold filtering, highlight the boundaries of objects well, but they are sensitive to noise and unstable to lighting variations. Thus, traditional image processing methods have a number of limitations that must be taken into account when choosing an approach for analyzing thermographic images (Deng et al., 2022).

Modern machine learning techniques, in particular deep neural networks, provide more accurate and flexible solutions. For instance, convolutional neural networks (CNNs) have been successfully employed for automatic defect classification. Object detection models like YOLO and Mask R-CNN not only detect defects but also segment their boundaries, as shown by Xin et al. (2024). Deep learning's advantages include high accuracy, automatic adaptation to varying shooting conditions, and consideration of complex spatial relationships within images. However, these methods require substantial computational resources and extensive training on large datasets. To evaluate the effectiveness of thermal imaging processing techniques, a comparative analysis is presented in Table 1, which outlines the key advantages and disadvantages of each method.

Table 1. Comparison of different theromgraphic image processing methods

Method	Advantages	Disadvantages
Gaussian filter	Noise removal, image smoothing	Loss of detail, blurring of borders
Median filter	Removes pulse noise, preserves boundaries	High computational complexity for large gaps.
Averaging filter	Easy to implement, reduced intensity variations	Blurring of edges, and loss of fine details.
The correlation method	High accuracy with known standards	Requires reference images, high computational complexity
Morphological operations	Defect structures are well distinguished	Depend on the parameters of the structural elements
Classical segmentation methods (Canny, threshold filtering)	Clearly define the boundaries of the objects.	Sensitive to noise, unstable to lighting variations

Analyzing the data presented in Table 1, we can see that traditional image processing techniques have several limitations, such as being sensitive to noise and requiring manual tuning of parameters. In contrast, deep learning algorithms are highly accurate at detecting defects and can automatically adjust to different shooting conditions and defect types. Deep neural networks like convolutional neural networks (CNNs) and object detection models (such as YOLO and Mask R-CNN) have demonstrated significant superiority in thermal image analysis tasks (Rizvi et al., 2024; Xie et al., 2023). These models not only detect defects but also classify them, reducing the chance of false positives and improving diagnostic accuracy.

It's worth noting that modern GPUs enable these models to be trained and deployed at high speeds, making them suitable for industrial use in real-world settings.

In addition, the use of deep learning allows us to reduce the impact of the human factor in the diagnostic process, which is particularly important for non-destructive testing. In traditional methods, the accuracy of defect detection depends heavily on the experience and expertise of the expert interpreting the images. Human errors can lead to incorrect assessments of the condition of pipes, potentially causing accidents or unnecessary repairs. By introducing automated systems based on neural networks, we can minimize subjective errors and ensure consistent analysis quality.

Despite the fact that deep learning methods require significant computing power and the availability of labeled data, their use is justified by the high efficiency of image processing. Moreover, in recent years, optimized versions of neural network architectures have appeared, such as MobileNet and EfficientDet, which allow machine learning to be applied even in resource-limited environments (Huang et al., 2024). This makes deep learning available for industrial applications, including automated quality control and pipeline health monitoring.

Special attention should be given to integrating various image processing techniques to improve diagnostic accuracy. Several studies suggest combining traditional pre-processing methods with deep learning techniques. For instance, filtering images before feeding them into a neural network can help improve the selection of defective contours, and using morphological operations can enhance segmentation accuracy. Therefore, hybrid methods that combine classical approaches with modern machine learning algorithms may produce the best results when analyzing thermal images.

Based on the analysis, we can conclude that deep learning methods significantly outperform traditional mathematical approaches in terms of accuracy, flexibility, and automation. Although they have a high computational cost, their use is justifiable due to their high efficiency in detecting defects.

In this regard, it would be advisable to use neural network models tailored to the specific needs of steel pipe analysis in order to develop software for processing thermographic images. The choice of an image processing technique should be based on a combination of factors, such as accuracy, complexity of calculations, adaptability, and the ability to interpret results automatically.

After analyzing the main approaches to processing thermographic images and identifying the priority direction for future work, we will examine the existing dataset and perform operations such as preprocessing, normalizing, cleaning, and labeling.

3. DATA SET FORMATION AND PREPROCESSING

In total, five different types of pipes were included in the data set that was studied. The pipes that carried seawater to the pumps were made of galvanized steel and were seamless, with outer diameters ranging from 377 to 219 millimeters and a wall thickness of 13 millimeters. The fresh water pipes that supplied the heat exchanger and the water pipes that led to the reducer and the oil pipes that went to the heat exchanger were all made of phosphated steel and were also seamless, with outside diameters varying from 114 to 89 millimeters and wall thicknesses ranging from 6 to 5.5 millimeters. All of the fresh water supply and discharge pipes had the same characteristics as those mentioned above, except for their standard size, which was an outside diameter of 89 millimeters with a wall thickness of 5.5 millimeters. Examples of some of the pipes that were studied can be seen in Figure 1.

Figure 1. Types of pipes under study.

To conduct the study, a dataset was created from 210 thermal imaging data collected on a cargo vessel, such as a self-discharging bulk carrier. The dataset was generated based on thermograms of the cooling system of a medium-sized, six-cylinder, four-stroke engine with a power output of 2,500 kilowatts (kW), captured using the GOYOJO handheld thermal imaging camera. Additionally, thermal and spatial information was provided for each image to facilitate data processing. The distribution of thermal imaging data by pipe group is shown in Figure 2.

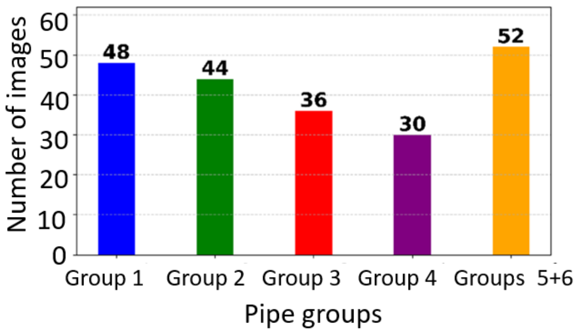

Figure 2. Distribution of the number of thermographic images by pipe groups

1 - Seawater supply pipes to the pumps;
2 - Fresh water supply pipes to the heat exchanger;
3 - Water supply pipes to the gearbox;
4 - Oil supply pipes to the heat exchanger;
5 - Fresh water intake pipes from the engine;
6 - Fresh water outlet pipes from the engine.

For further analysis and machine learning model training, it is essential to categorize the defects identified in the thermal images. In this research, three primary classes of defects were identified: point, extended, and volumetric.
The marking of thermographic images is done using the instance segmentation method. This is a computer vision technique that allows us not only to identify areas of objects in an image, but also to differentiate them from each other by assigning each object a unique label.
Unlike other types of segmentation, like semantic segmentation, which groups all objects of the same class into a single mask, instance segmentation marks each object individually. This is particularly useful for defect detection and analysis (Yılmaz et al., 2025).
In this study, we used specialized cvat software to perform the markup.
Each image was carefully marked up, taking into account the contours of the identified defects. This involved careful selection of the boundaries of the affected areas and assigning each defect a class based on the classification system. In each image, each defect was given a corresponding class label, which allowed for automatic differentiation between different types of defects during further processing. Examples of thermographic image marking are shown in Figure 3.

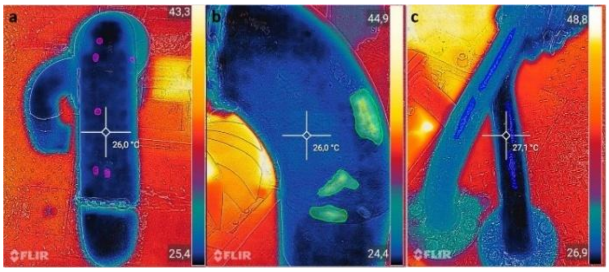

Figure 3. Examples of marking defects on thermograms. a – Point defects; b – volume defects; c – Extended defects.

To ensure compatibility with the neural network architecture used, the annotated data was converted to a format supported by the YOLO model with segmentation capability. Unlike the classical object detection problem, which uses bounding boxes to represent annotations, the instance segmentation problem requires a more precise description of the shape of an object.
The study used YOLO Segment, a modified annotation structure, which includes not only the object class but also vector coordinates describing the contour of the object (mask). The original markup was made in COCO format using CVAT and contained segmentation in the form of polygonal contours.
To convert the markup to the YOLO Segment format, we implemented a proprietary converter script that automatically extracts information from COCO annotations and converts mask coordinates to normalized values. The script then generates corresponding .txt files in the YOLO format. Each annotation file contained a line-by-line description of all the objects in the image, in the format of <class_id> x_1 y_1 x_2 y_2 ... x_n y_n. Here, x_1, y_1, ..., x_n and y_n are the normalized coordinates of the points that form the outline of the object's mask. This format meets the requirements of the YOLO algorithm for segmenting instances with high accuracy.
After the conversion process was complete, the dataset was organized according to the requirements of the YOLO model. The images and their annotations were divided into three subsets: training, validation, and testing.
To better understand the results of thermographic data processing, we conducted an analysis based on statistical and visual features of the collected dataset. This allowed us to identify patterns in the distribution of defects across images, evaluate the balance of classes, and determine the level of annotation saturation in images. Figure 4 shows how many images have at least one instance of each defect class.

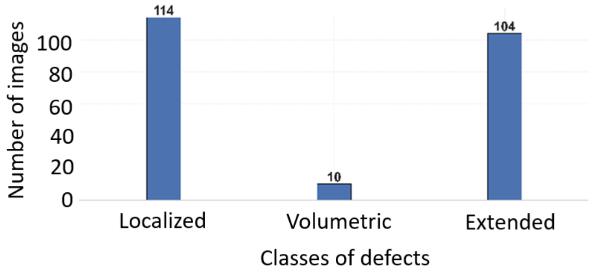

Figure 4. The number of images containing objects of each class

This allows you to assess how evenly the classes are represented in the dataset. The graph shows that the majority of images contain Localized defects, which may indicate a high prevalence of this type of damage, as well as a possible imbalance in the original dataset. Since the Volumetric class is less common, it may be necessary to use balancing techniques when training the model, such as class weighting or generating synthetic examples.

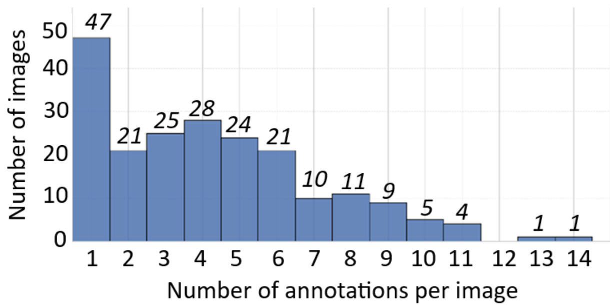

Figure 5. Distribution of the number of annotations across images

Figure 5 shows the distribution of the number of annotated objects (defects) in each image. This visualization is important for understanding the complexity of the images presented to the analysis algorithm. Most images contain between 1 and 6 objects, which is typical for simple cases. However, there are also images with 10 or more objects, indicating more complex cases with multiple defects. These cases are especially challenging when the damage is localized and extensive. When training and testing a model, it is essential to consider the variety of object densities in order to make it robust to both simple and complex cases that require high accuracy in segmentation. The data used to train the model served as the basis for selecting the model architecture and learning approaches, including techniques for balancing classes and generating synthetic examples.. Thus, statistical and visual analysis of the dataset is an essential stage in building a system for automatic defect detection in thermographic images, providing a solid foundation for subsequent engineering implementation.

4. TRAINING A NEURAL NETWORK MODEL TO IDENTIFY PIPE DEFECTS IN THERMOGRAPHIC IMAGES

The main purpose of the training was to create a system that could accurately determine both the presence of defects in steel pipes from thermographic images and their exact contours, thereby providing a full-fledged visual analysis. The study used the YOLOv8n-seg modification, which is compact (about 3.25 million parameters) and relatively low computational complexity (about 12 GFLOPs), which makes it suitable for systems with limited computing resources (Huang et al., 2024).
Figure 6 shows the architectural diagram of the YOLOv8s-Seg model, which reflects the key principles behind the construction of the YOLOv8 segmentation variant. This diagram illustrates the RepBlock structure, the use of SimConv convolution layers in the "neck" part, as well as the cascaded feature aggregation that ensures the effectiveness of predictions while minimizing resource consumption. This architecture allows for high inference speeds while maintaining accuracy comparable to that of more complex models.
In the context of detecting pipe defects in thermographic images, this model has proven to be particularly effective. It provides detailed and accurate detection of damage areas, while maintaining high performance. The neural network was trained using the Ultralytics open-source library, which offers a high-level API for working with YOLOv8 models.

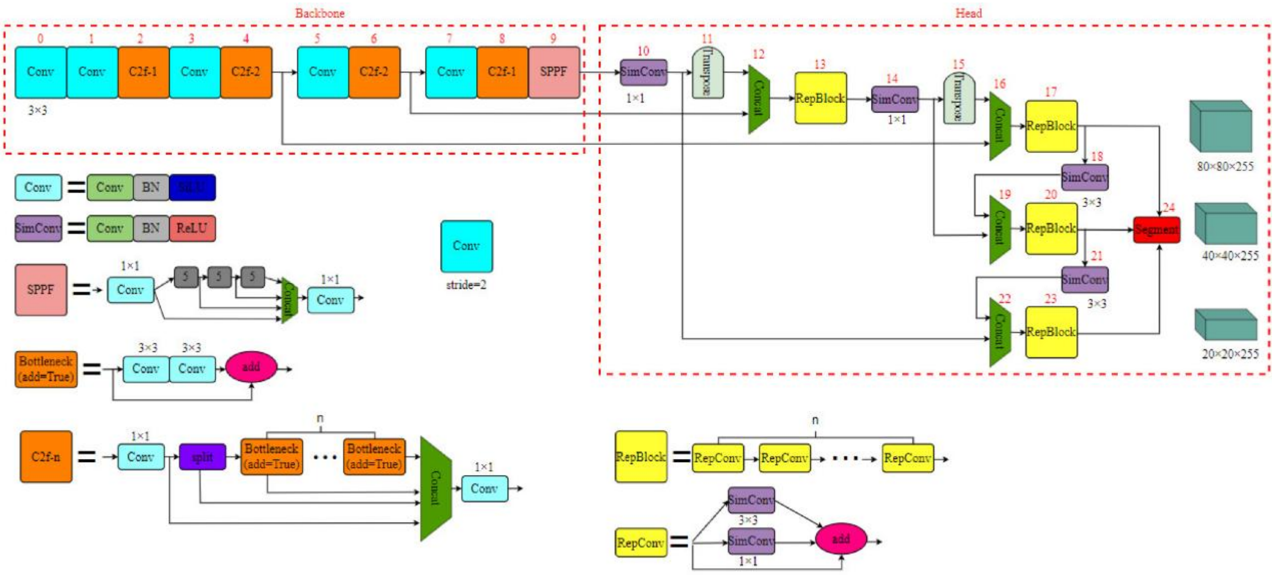

Figure 6. YOLOv8-Seg neural network architecture (Anupama et al., 2024)

At the first stage, the YOLO class was imported from the ultralytics library using the command 'from ultralytics import YOLO'. Then, the YOLOv8 segmentation model was loaded, which had been pre-trained on a large dataset (COCO or a similar dataset). In this study, a lightweight version of YOLOv8n-seg was chosen, which provided fast operation with acceptable accuracy.
Figure 7 shows the neural network architecture. During the training process, images from the training set were fed into the neural network, which generated predictions of defect masks. These predictions were then compared with the true annotations, and the loss function was calculated based on the differences between the two. The loss function allowed the model's weights to be adjusted using the backpropagation method.
The overall loss function was composed of several components and was represented by the following formula:

$$L_{total} = \lambda_{box} * L_{box} + \lambda_{cls} * L_{cls} + \lambda_{seg} * L_{seg} \quad (1)$$

where: L_{box}- localization error; L_{cls}- classification error; L_{seg}- seg- segmentation error; $\lambda_{box}, \lambda_{cls}, \lambda_{seg}$- Weighting factors that determine the contribution of each component to the overall loss function.

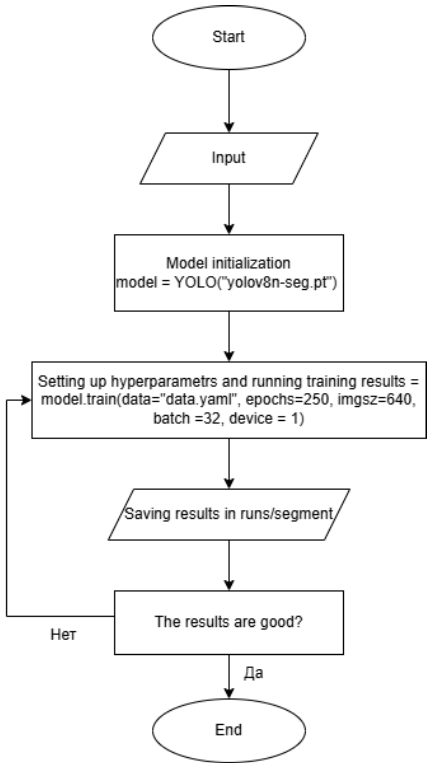

Figure 7. Neural network learning algorithm

To evaluate the quality of the predicted bounding boxes, the IoU Loss function is used, which measures the discrepancy between the predicted frame and the true annotation. One of the popular implementations is Complete IoU (CIoU) Loss, which takes into account not only the intersection area, but also the distance between the frame centers and the aspect ratio.:

$$L_{box} = 1 - IOU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \qquad (2)$$

where IoU is the Intersection over Union; *p* is the Euclidean distance between the centers of the frames; *c* is the diagonal of the minimum enclosing rectangle that minimally covers both the predicted and the true frame; *u* is a measure of the difference in aspect ratio between the predicted frame and the true frame; *α* is the parameter that regulates the contribution different components in the overall loss function. It depends on the values of IoU and u and is calculated as follows (3).

$$\alpha = \frac{1}{1 + IoU + v} \qquad (3)$$

The predictions of the model regarding the belonging of an object to a particular class are evaluated using Binary Cross-Entropy (BCE).

$$L_{cls} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(p_i) + (1 - y_i)\log(1 - p_i)] \quad (4)$$

where $y_i \in \{0,1\}$ - the true class label, $p_i \in [0,1]$ - predicted probability, N- number of examples.
The quality of object segmentation, that is, the prediction of the shape of defects, is also measured using BCE Loss, since the task is presented as a binary mask on each pixel using the formula:

$$L_{seg} = -\frac{1}{HW}\sum_{i=1}^{H}\sum_{j=1}^{W}[m_{ij}\log(\hat{m}_{ij}) + (1 - m_{ij})\log(1 - \hat{m}_{ij})] \qquad (5)$$

Where m_{ij} is the true value of the mask in pixels (i, j), $\hat{m}_{ij}$ is the predicted probability of an object at pixel (i, j), H and W are the height and width of the mask, respectively.
At each step, the backpropagation process occurred, ensuring consistent adjustments to the model's parameters. After each iteration, the model was validated on a separate sample, allowing for the timely identification of potential signs of overfitting (Hussain et al., 2023). Training was conducted on an NVIDIA GeForce RTX 2080 Ti graphics card with 11GB019 MB of video memory, which provided high performance in processing images of size 640x640 pixels, accelerating the learning process within an acceptable timeframe. Using a GPU enabled simultaneous processing of batches of images, reducing overall training time.
After setting up the basic parameters for the model and starting the main learning process, we conducted a series of experiments to optimize the quality of the defect segmentation. At this stage, one of the key tasks was to select suitable hyperparameters that would influence the final performance of the model. These parameters included batch size, learning rate, number of training epochs, data augmentation methods, and optimizer. We selected these parameters empirically and then analyzed the results obtained from validation and testing subsets.
One of the experiments we conducted was training the YOLOv8-seg model without using any specialized data augmentation techniques. The configuration for this experiment included 250 training epochs, a batch size of 16, an image size of 640x640 pixels, the AdamW optimizer, and a learning rate of 0.01.

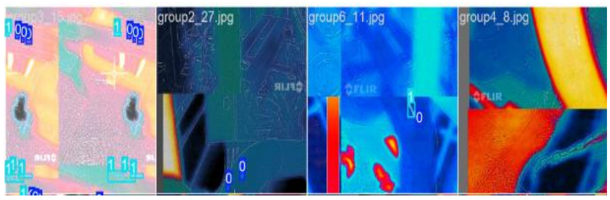

Figure 8. Example of images with annotations used as input for a neural network.

During model training, a batch of images and their corresponding annotations (masks) is provided at each step. The neural network is then trained consistently to extract features and make predictions from these images. Figure 8 shows an example of such a batch being input to a neural network. This allows us to visually evaluate the variety of images and the shapes of the marked objects (Sadyraliev et al., 2025).

At the end of the training process, the metrics were analyzed on a test sample. To evaluate the quality of the computer vision model, we used the following metrics: accuracy (precision), completeness (recall), and mean average precision (mAP). (Shcherban et al., 2024)

During the first iteration, the results of segmentation quality were not satisfactory due to the presence of a large number of point defects and image noise. We decided to use the copy-paste augmentation method to improve segmentation quality and increase model stability. This will help us to cope with class imbalance. (Mohammadzadeh et al., 2025)

During this procedure, we used the same parameters as in the baseline case: 250 epochs, 16 batch size, image resolution of 640x640 pixels, AdamW optimizer, and an initial learning rate of 0.01. The only difference was the activation of the additional Copy-Paste data augmentation technique. Figure 9 shows a sample batch of images with Copy-Pasted content.

Visual analysis of the metric graphs suggests more stable model performance and improved mean Average Precision (mAP) values. Figure 10 displays the dynamics of the metrics, including mAP@0.5 and mAP@0.5:0.95, for frames and masks during the learning process using Copy-Paste.

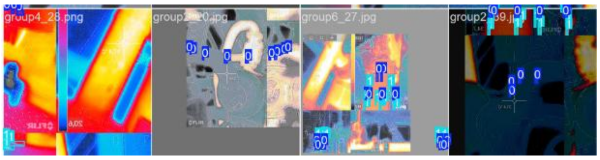

Figure 9. Example of a training batch using Copy-Paste augmentation

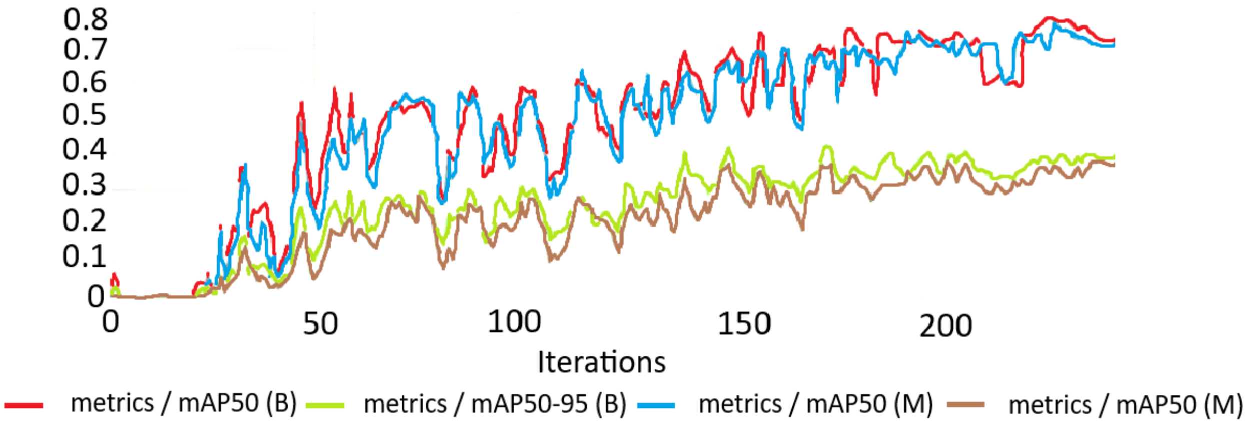

Figure 10. Graphs of mAP metrics for bounding boxes and masks (Copy-Paste)

A summary of the results for all defects is presented in Table 2 and Table 3.

Table 2 – Learning outcomes with Copy-Paste (for Box)

Type of defect	Box Precision	Box Recall	Box mAP@50	Box mAP@50-95
Localized	0.537	0.429	0.413	0.172
Extended	0.474	0.800	0.766	0.538
Volumetric	0.920	0.870	0.995	0.522
Average	0.671	0.726	0.725	0.411

Table 3 - Results of Training with Copy-Paste (For Mask)

Type of defect	Mask Precision	Mask Recall	Mask mAP@50	Mask mAP@50-95
Localized	0.519	0.286	0.385	0.176
Extended	0.586	0.802	0.816	0.392
Volumetric	0.922	0.871	0.995	0.564
Average	0.695	0.652	0.732	0.377

An analysis of the training results of the YOLOv8n-seg model with Copy-Paste augmentation reveals that the best performance was achieved for volumetric defects. The precision was 0.922, and the recall was 0.871. This indicates the high reliability of the model for detecting and segmenting large defects. For extended defects, good indicators were obtained as well (precision: 0.586; recall: 0.802). However, the accuracy is lower here, which may be attributed to the presence of false positives or the difficulty in defining the boundaries of these defects.

On the other hand, the model demonstrated relatively poor results for point defects. The precision for localized defects was 0.519 and the recall 0.286. This can be explained by the small size of these defects, image noise, and imbalanced classes in the dataset.; The average values of the metrics (mAP@50 and mAP@50-95) indicate that the model performs better on large and extended defects compared to point defects.

The use of Copy-Paste augmentation improves the quality of the segmentation, especially for classes with insufficient data. However, to further improve the accuracy, especially when dealing with small defects, it is recommended to explore additional augmentation methods and increase the size of the dataset. This will allow for more stable results across all types of defects in the future (Shcherban et al., 2023).

Graphs of mAP metrics for bounding boxes and masks (Copy-Paste) is presented on Figure 11.

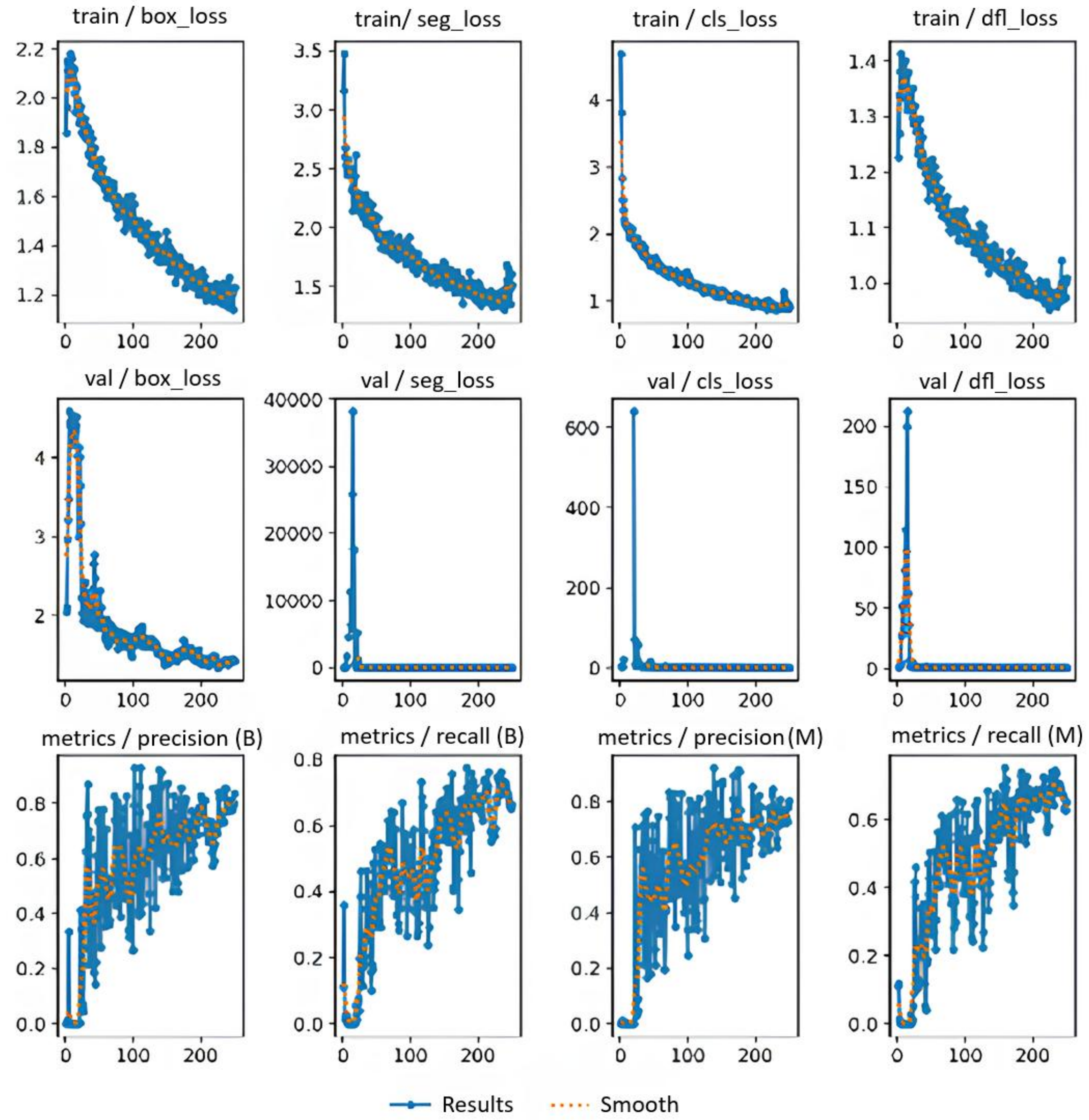

Figure 11. Graphs of mAP metrics for bounding boxes and masks (Copy-Paste)

Figure 12 demonstrates examples of predictions from a test dataset that was trained using Copy-Paste.

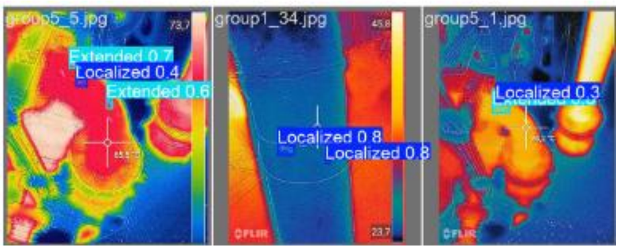

Figure 12. Examples of model predictions on a test set using Copy-Paste

The model has demonstrated improved accuracy in processing small objects and complex scenes with multiple defects, confirming the effectiveness of the augmentation technique.

The use of the Copy-Paste augmentation technique has made it possible to achieve steady improvement in segmentation quality, especially in challenging cases where gaps or blurry masks were observed in the initial experiment. The method proved to be effective when working with a limited number of images and can be recommended for industrial segmentation tasks with similar data characteristics.

After training the YOLOv8n-seg neural network for instance segmentation of defects in thermal images, stable and repeatable results were obtained that ensure

high-quality object segmentation of various complexities. The model showed high accuracy and generalization ability with a small training sample size, particularly after introducing the Copy-Paste method.
The experiments allowed us to identify the impact of key hyperparameters and preprocessing techniques on the final performance of the model. In the first basic experiment, we achieved acceptable prediction accuracy, but we encountered difficulties in detecting minor defects. However, in the second experiment, after adding Copy-Paste, we were able to significantly improve the mean average precision (mAP) and increase accuracy for the Localized class. This suggests the usefulness of using synthetic augmentation techniques in situations with limited data availability.

5. CONCLUSION

The studies conducted have shown the effectiveness of using thermodiagnostics to generate data on the technical condition of pipes in cooling systems. A neural network was created, which allows for the detection of various types of pipe defects (point, extended, and volumetric), based on thermographic images. The use of restrictive masks to detect defects, as well as the instance segmentation and copy-paste augmentation techniques, have been confirmed to improve the quality of the image processing results by the trained neural network.
At the same time, the following values were obtained for various types of defects according to key metrics: The neural network model demonstrated high accuracy in detecting volumetric defects (Precision = 0.922, Recall = 0.871). It also showed good results for extended defects (Precision = 0.586, Recall = 0.802), but relatively low rates for point defects (Precision = 0.519, Recall = 0.286).. This difference in efficiency can be explained by the characteristics of each type of defect. Large volume damages are easily distinguishable on thermograms, whereas small point defects often get masked by noise and require more detailed analysis.
The results show that the neural network is able to successfully segment and classify the main types of defects in steel pipes on thermographic images. However, it needs further development in order to improve its performance when dealing with smaller objects.
To improve the quality of the model, it is recommended to increase the size of the training dataset, especially for small point defects. Additionally, using data augmentation techniques could also help improve the model's performance.
In the future, it will be necessary to determine how the accuracy of detecting and identifying different types of defects varies depending on the thickness of the pipe wall, the temperature of the liquid being pumped, and the temperature of the metal surface of the pipe (the relationship between these factors and the ability of the neural network to detect and categorize defects).
The work on establishing the quality standards for thermograms for further processing remains an unresolved issue. This requires setting the minimum and maximum focal length for the camera shooting in the infrared range, as well as the tube under study. Additionally, it is important to determine the permissible angles of the camera relative to the tube. In particular, we need to understand how the placement of the camera affects the quality of the data set generated and the possibility of neural network processing.
Another promising area of research is the development of software that allows for the processing of thermographic images of pipes and the detection of defects. Such software is currently being developed. It should not only be able to record the presence or absence of a defect and categorize it, but also determine its acceptability. This, in turn, requires the creation of a database of regulatory requirements for different types of pipes.
However, it should be noted that the thermographic method and neural network-based defect recognition may not be effective in certain cases. Specifically, when examining pipes that transport liquids with temperatures similar to the ambient environment, as well as when detecting defects in pipe joints. Nevertheless, the results of the study allow for a significant expansion and simplification of the procedures for thermographic diagnostics of steel pipes, and can improve the safety of systems that transport aggressive and high-temperature liquids.
The developed approach also has practical potential for industrial inspection. By reducing the need for manual analysis of thermographic images, the system could make the inspection process faster and provide more consistent results. Its application could be particularly useful as a supporting tool for maintenance personnel, helping them identify potentially dangerous defects at an early stage. However, before practical implementation, further testing with a larger and more diverse dataset is needed, especially for small point defects and under different operating and imaging conditions.

References:

Alshawabkeh, S., Wu, L., Dong, D., Cheng, Y., & Li, L. (2025). A Hybrid Approach for Pavement Crack Detection Using Mask R-CNN and Vision Transformer Model. *Computers, Materials & Continua, 82*(1).

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., ... & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data, 8*, 1-74.

Anupama, M. A., Chhabra, K., Ghosh, A., & Thavva, R. S. R. (2024, January). Comparative Analysis of Deep Learning Models for Car Part Image Segmentation. In International Conference on Data Management, Analytics & Innovation (pp. 267-279). Singapore: Springer Nature Singapore.

Ciampiconi, L., Elwood, A., Leonardi, M., Mohamed, A., & Rozza, A. (2023). A survey and taxonomy of loss functions in machine learning. arXiv preprint arXiv:2301.05579.

Deng, J., Singh, A., Zhou, Y., Lu, Y., & Lee, V. C. S. (2022). Review on computer vision-based crack detection and quantification methodologies for civil structures. *Construction and Building Materials, 356*, 129238.

Golodov, V. A., & Maltseva, A. A. (2022). Approach to weld segmentation and defect classification in radiographic images of pipe welds. *Ndt & E International, 127*, 102597.

Han, S., Yang, F., Jiang, H., Yang, G., Wang, D., & Zhang, N. (2022). Statistical analysis of infrared thermogram for CNN-based electrical equipment identification methods. *Applied Artificial Intelligence, 36*(1), 2004348.

Huang, J., Zhang, X., Jia, L., & Zhou, Y. (2024). A high-speed YOLO detection model for steel surface defects with the channel residual convolution and fusion-distribution. *Measurement Science and Technology, 35*(10), 105410.

Huang, X., Han, P., Fu, Y., & Zhang, Y. (2024, April). YOLOv8-BVC: Lightweight Industrial Defect Detection. In 2024 5th International Conference on Computer Vision, Image and Deep Learning (CVIDL) (pp. 1220-1226). IEEE.

Hussain, M. (2023). YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection. *Machines, 11*(7), 677.

Kim, H. J., Shrestha, A., Sapkota, E., Pokharel, A., Pandey, S., Kim, C. S., & Shrestha, R. (2022). A study on the effectiveness of spatial filters on thermal image pre-processing and correlation technique for quantifying defect size. *Sensors, 22*(22), 8965.

Liu, G., Gao, W., Liu, W., Chen, Y., Wang, T., Xie, Y., ... & Li, Z. (2024). Automatic defect classification for infrared thermography in CFRP based on deep learning dense convolutional neural network. *Journal of Nondestructive Evaluation, 43*(3), 73.

Liu, Z., Lu, G., Liu, X., Jiang, X., & Lodewijks, G. (2017). Image processing algorithms for crack detection in welded structures via pulsed eddy current thermal imaging. *IEEE Instrumentation & Measurement Magazine, 20*(4), 34-44.

Ma, Q., Tian, G., Zeng, Y., Li, R., Song, H., Wang, Z., ... & Zeng, K. (2021). Pipeline in-line inspection method, instrumentation and data management. *Sensors, 21*(11), 3862.

Mohammadzadeh, M., Günay, E. E., Jackman, J., Kremer, G. E., & Kremer, P. (2025). Utilization of Data Augmentation Techniques in Automated Inspection Systems for Defect Detection in Metals With Limited Data. *Journal of Advanced Manufacturing and Processing*, e70011.

Qureshi, U. R., Rashid, A., Altini, N., Bevilacqua, V., & La Scala, M. (2024). Radiometric Infrared Thermography of Solar Photovoltaic Systems: An Explainable Predictive Maintenance Approach for Remote Aerial Diagnostic Monitoring. *Smart Cities, 7*(3), 1261-1288.

Rizvi, S. Z., Jamil, M., & Huang, W. (2024). Enhanced defect detection on wind turbine blades using binary segmentation masks and YOLO. *Computers and Electrical Engineering, 120*, 109615.

Sadyraliev, Z., Zhumanbayeva, T. K., Dybina, O. V., Krutilin, A. A., & Amirkhanyan, G. M. (2025). Big Data and Artificial Intelligence In Information Systems Of Product Quality Management. *Proceedings on Engineering Sciences, 7*(1), 119-128.

Shcherban, P., & Abu-Khamdi, R. (2024). Statistical Study on Domestic Gas Boiler Failures Using Various Software Platforms. *Ingeniería e Investigación, 44*(3), e112017.

Shcherban, P., Sokolov, A., Reda Abu H. (2023) Study of failure statistics of Cavitators in the fuel oil facilities Through the application of regression And cluster analysis, *Proceedings on Engineering Sciences, 5*(1), 39-48.

Ullah, I., Khan, R. U., Yang, F., & Wuttisittikulkij, L. (2020). Deep learning image-based defect detection in high voltage electrical equipment. *Energies, 13*(2), 392.

Xie, Y., Hu, W., Xie, S., & He, L. (2023). Surface defect detection algorithm based on feature-enhanced YOLO. *Cognitive Computation, 15*(2), 565-579.

Xin, M., Xu, C., Gao, J., Wang, Y., & Wang, B. (2024). High-Precision Recognition Algorithm for Equipment Defects Based on Mask R-CNN Algorithm Framework in Power System. *Processes, 12*(12), 2940.

Yılmaz, Y., Nayır, S., & Erdoğdu, Ş. (2025). Real-time detection of steel corrosion defects using semantic and instance segmentation models based on deep learning. *Materials Today Communications, 44*, 112050.

Shcherban Pavel Sergeyevich
Immanuel Kant Baltic Federal University, Branch Scientific Cluster Institute of High Technologies
Kaliningrad, Russia
ursa-maior@yandex.ru
ORCID: 0000-0001-5106-7852

Mazur Ekaterina Vladimirovna
Department of Shipbuilding, Ship Repair and Marine Engineering, Kaliningrad State Technical University, Russia
Kaliningrad, Russia
ekaterina.mazur@gmail.com
ORCID: 0000-0003-2657-6369

Eranosian Sos Samvelovich
Yerevan State University, Yerevan, Armenia
alikos20001@gmail.com
ORCID: 0009-0000-3573-6411

Fedor V. Piskun
Immanuel Kant Baltic Federal University, Branch Scientific Cluster Institute of High Technologies
Kaliningrad, Russia