Research on Engineering Structures & Materials
www.jresm.org
Research Group

Research Article

Detection of defects in steel oil and gas operating pipelines using a neural network, based on the results of optical diagnostics

Shcherban Pavel Sergeyevich [*,1,a], Potapov Boris Vasilievich [2, 3b], Karagadyan Artur Nairievich[1,4c]

[1]*Baltic Federal University of Immanuel Kant, BSC "Institute of high technologies", Kaliningrad, Russia*
[2] *RTU MIREA Institute of Information Technology*
[3] *Russian State Agrarian University - Moscow Timiryzev Agricultural Academy*
[4] *JSC VNIIST, Moscow, Russia*

Article Info

Article History:

Received 01 Feb 2025

Accepted 15 June 2025

Keywords:

Metal pipes;
Optical control;
Neural network;
Surface diagnostics;
Pipe defects

Abstract

In today's world, steel pipelines are an essential part of the infrastructure for various industries, such as oil and gas, energy, and industry. However, their flaws can lead to major accidents, resource loss, and environmental hazards. Therefore, it is crucial to identify defects early on and accurately determine them to ensure the safe and efficient operation of pipeline transportation systems. During research, a technique was developed for automatically detecting and classifying defects in steel oil and gas pipelines using neural networks. This method processes image data from an in-line inspection system. The initial data for this analysis comes from optical inspections of oil and gas pipelines. The study involved developing and testing a method that can detect defects in oil pipelines automatically with significant accuracy using data from optical examinations. Optical examination can be performed using both specialized endoscopic equipment and a crawler robot. The processing of diagnostic results using a neural network and data library can be done on an external medium or integrated into a diagnostic system for in-line monitoring. Testing of the pipe defect recognition method during optical inspection of the inner pipe surface demonstrated good results for detecting corrosion and erosion defects (response rates of 67% and 68%, respectively, with mAP values of up to 49% and 65%). Crack-type defects were identified with lower accuracy (response rate of 57%, mAP up to 42%).

1. Introduction

The existing methods of inline diagnostics of steel pipelines include the use of ultrasonic, eddy current, and magnetic flaw detectors. The use of this devices allows us to determine the type of defect and its geometric parameters with high accuracy. However, in steel pipes with small diameters, such as those used in oil and gas fields systems with diameters less than 250 millimeters, it can be challenging to perform these types of diagnostic procedures due to the lack of specialized equipment and technical tools [1].

For pipes of this type, an optical examination of the inner surface is a simpler and more natural method, according to which one can judge the integrity and reliability of the pipes. The length of these pipes can be considerable, amounting to several kilometers, significantly complicating the processing of photographic and video data obtained during the examination. This process requires

*Corresponding author: ursa-maior@yandex.ru
[a]orcid.org/0000-0001-5106-7852; [b]orcid.org/0000-0002-9635-3809; [c]orcid.org/0000-0002-8427-8362
DOI: http://dx.doi.org/10.17515/resm2025-650st0201rs

a lot of time, but an increase in processing speed and data quality can be achieved through the use of a specialized neural network. Optical control methods in conjunction with neural networks for processing and detection of emergency situations and defects, are developing extremely actively and are used in various industrial and transport sectors [2]. The efficiency, detection speed, and accuracy of neural networks in identifying individual defects and emergency situations are steadily improving.

In recent scientific works [3], there has been an increase in the use of various machine learning methods for diagnosing oil and gas pipes and identifying internal defects. In that regard, the most interesting studies are [4,5], which conduct a deep analysis of the applied use of various machine learning algorithms and establish the possibility of detecting and segmenting defects in oil and gas pipelines in the optical range (in particular, corrosion and crack defects). In addition, research [6,7] has significant potential, where corrosion and erosion datasets are formed and defects of this type in pipelines are determined.

Researchers are developing and testing various algorithms [8, 9], which forms a representative array of information on their effectiveness for solving specific production problems. This confirms that the use of computer vision methods will soon find a direct response in the methodological foundations of diagnostics of oil and gas equipment in general and determining pipe defects in particular. The development of a general methodology and further adoption of the corresponding standard at the national and international level are expected. Obviously, this will happen based on the results of testing various algorithms, taking into account their effectiveness and accuracy of detection and segmentation of defects. It is also likely that the most effective approaches will be determined for various control methods (optical, ultrasonic, eddy current, radiographic).

The aim of this research was to develop a neural network capable of processing and segmenting defects with high accuracy, including bulges, corrosion, cracks, and erosion, in oil and gas steel pipes during optical examination. The distinctive features of this study from any previously published works in scientific periodicals is that to form a dataset of pipes with defects, JSC All-Union Scientific Research Institute for the Construction and Operation of Pipelines and Fuel and Energy Complex Facilities (VNIIST) used video and photo materials from optical inspection of field oil and gas pipelines collected by the institute in the period of 1980-2020s (the institute itself was founded in the USSR in 1948). In addition, based on the dataset, a specialized YOLO (You Only Look Once) neural network was trained to recognize specific defects in oil and gas equipment. (The main abbreviations are presented in Appendix 1.).

During the research, we analyzed the most critical types of defects and created a dataset for each one (based on internal surveys of steel pipes, used in oil and gas field transportation systems). For each type of defect, the dataset contained at least 100 images. Then, we trained a neural network to segment images of bulges, corrosion, cracks, and erosion in steel pipes using the results of optical examination. Based on this neural network, the software was developed. The results of the study will lead to the creation of an improved methodological apparatus and software solutions for robotic means of conducting diagnostic examination of pipeline systems. This is an important part of the technical condition monitoring system, which is necessary for a long-term and safe operation of objects.

The remainder of this paper is organized as follows: the section “In-line diagnostics and the main types of defects in steel pipelines” provides an overview of approaches for gathering of optical data from the inner pipe surfaces. Section “The formation of a neural network and the creation of a library of defects” outlines dataset construction and preprocessing. Also it contains information about the selection and architecture of neural network to be used with the data. In section “Training a neural network to detect pipe defects and analyzing the results” we provide both visual and quantitative results. These include images of recognized defects and metric plots—to evaluate the model’s performance. We conclude the paper in section “Conclusions” where we summarize the obtained information and outline future researches.

2. In-Line Diagnostics and The Main Types of Defects in Steel Pipelines

The formation of a database of defects was based on images obtained from inspections of oil and gas systems transporting borehole products. These images were collected through endoscopic examinations of the systems or using a crawler robot (see Figure 1). All sites where optical data was collected were made from steel pipes. Among the most common defects in the received information package are bulges, corrosion, cracks, and erosion. This is generally due to a wide range of conditions (physical and chemical processes) that change and predetermine these defects' occurrence. Such violations lead to a significant decrease in the reliability and strength characteristics of the pipe system, resulting in the pipes ceasing to fulfill their main transport function [10].

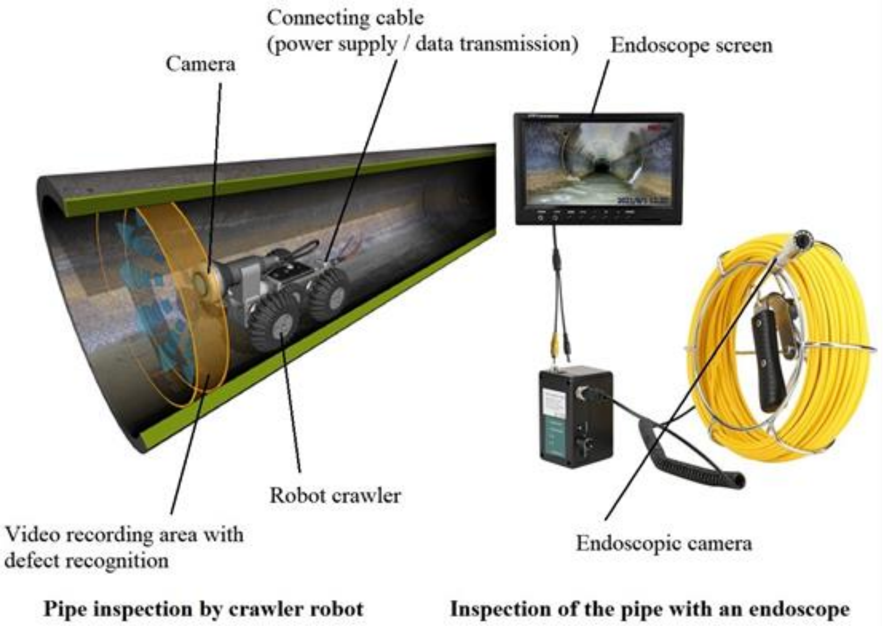

Fig. 1. Examples of a crawler and an in-tube endoscope for optical examination of pipes

The issue of the frequency of frame capture of the inner surface of the pipes is crucial for ongoing research. Therefore, it seems reasonable to split the video stream into frames at a frequency of 20 frames per second for optical diagnostics. Based on the speed of movement of a self-propelled robotic platform or endoscope moving in a straight line, this frequency is the minimum necessary. When diagnostic equipment moves at a speed of 1 cm per second, this frame rate allows us to capture damages up to 0.5 mm in size. If smaller defects need to be captured, the frame rate needs to be increased. The recorded video stream needs to be processed and analyzed frame by frame. After that, the results of processing are analyzed and combined every 20 consecutive frames. This process allows us to identify the most damaged areas, their approximate location, size, geometric shape, and color distribution.

To create an accurate image of the pipe's surface, a wide-angle camera and illuminator system should be used. They should be fixed in such a way that the entire pipe falls within the camera's field of view. Then, the part of the frame around the perimeter of the pipe that is in focus is selected, forming a ring. The center of this ring is the optical center of the lens [11]. This process results in an annular panoramic image of the pipe. Next, either direct processing of the video data can be done, or each individual frame can be preprocessed to reduce geometric distortion. Geometric distortions can occur due to the lens's distortion. These distortions cause straight lines in the image to bend, creating a "barrel" effect that reduces distance from the optical center. However, the part of the image that corresponds to the plane of the optical axis remains undistorted.

No additional image preprocessing unit has been developed in the software that was created to eliminate distortion effects. Instead, a direct assessment of defects observed in the optics was carried out without scanning images. This approach has a slightly lower accuracy in general, particularly with respect to minor defects such as pitting and hairlines. However, it is also less resource-intensive and significantly faster.

There are numerous different types of defects on the inner surface of steel pipes, and their systematics and specific features are reflected in various regulatory documents. These include ISO 11971:2020 Steel and iron castings — Visual testing of surface quality, ISO 3183:2019 Petroleum and natural gas industries — Steel pipe for pipeline transportation systems; ISO 11960:2020 Petroleum and natural gas industries — Steel pipes for use as casing or tubing for wells; ISO 7005-1:2011 Part 1: Steel flanges for industrial and general service piping systems.

In their research, the authors [12,13] and [14] rely on international regulatory documents and the classification of defects. It should be noted that national or industry classifiers can also be used, where the characteristics of defects can be presented in more detail. National classifiers can be linked to pipe manufacturers and manufacturers of equipment used in diagnostics. Among Russian regulatory documents, the taxonomy of in-pipe defects is most fully presented in GOST R 59496-2021 Welded steel pipes. Welded joint defects. Terms and definitions and in the current OST 14-82-82 "Industry quality management system for ferrous metallurgy products. Departmental quality control of products. Seamless rolled steel pipes. Surface defects. Terms and definitions". Unfortunately, at the moment, various scientific works on the detection of defects in steel pipes using neural network technologies do not cover all types of defects (regardless of their classification). This is often due to the lack of a representative and high-quality sample of defects for use in neural network training [15]. The defects can be divided into three categories: pipe geometry defects, pipe wall defects, and welded joint defects (Figure 2).

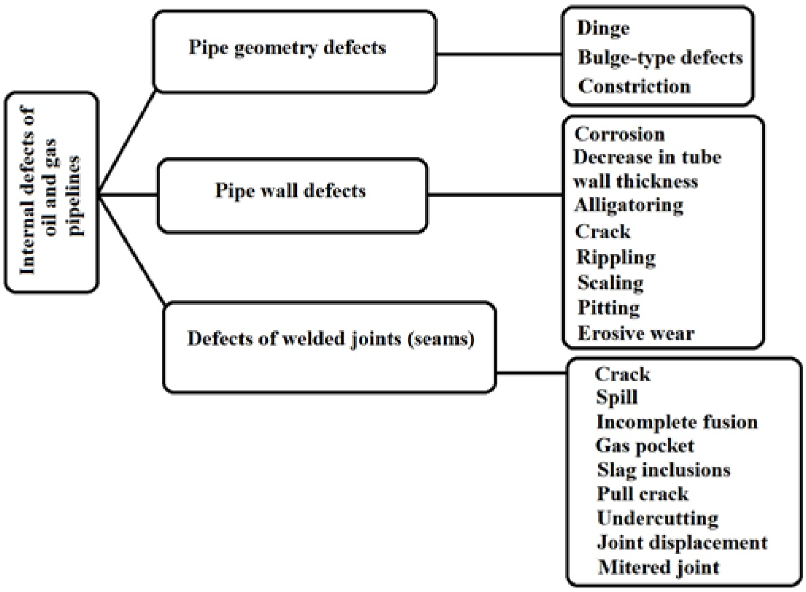

Fig. 2. Main categories of defect types in steel pipelines

Due to the lack of frame-by-frame pre-processing in this study, geometric distortions were not eliminated and images obtained from the side surface of the crawler or endoscope were not processed (only images from the front camera were used). As a result, some potential defects at pipe connections may have been missed. To solve this problem, it is necessary to consider the issue of distortions or, even better, conduct panoramic or side-by-side shooting. In such a case, defect detection will be maximized within the optical range [16].

In this regard, as part of the ongoing research, it was decided not to consider defects in pipe connections separately (since either a different type of survey or additional data preprocessing is required), but to limit ourselves to determining defects in the inner surface of the pipe body. The most common defects affecting the strength of the inner surface of the steel pipe body are the following: bulges, corrosion, cracks and erosion. Potholes and delaminations are relatively common, but there were no defects of these types in the available set of photo and video data used for the study. Let's take a closer look at the defects analyzed in the study.

The bulge is a local, gentle deflection of the pipe wall, with possible thinning in some areas. It indicates a local loss of strength and bearing capacity of the pipe, as well as a violation of the geometric shape. This may be caused by cyclic loads. Bulges can be convex or concave, and their presence leads to a change in the flow rate of liquid, which increases the impact of the transported material on the pipe walls in that area. It should be noted that the number of bulges is relatively small in the data set. This defect is rare and difficult to detect visually from inside the pipe. For steel pipes, the most common type of defect is corrosion. Corrosion can be of various types, such as rivulet, pitting, and solid. The generated data library takes into account all these types of corrosion damage to pipes. In general, corrosion refers to the spontaneous chemical interaction of metal with its environment, resulting in a change in its composition and properties. This process occurs without the need for external energy input. The corrosion process involves the chemical reaction between the metal and the medium.

Therefore, if a metal structure loses its bearing capacity due to exposure to a dry hydrocarbon gas stream containing abrasive particles, this is not considered as corrosion-related damage. If aggressive components such as moisture, hydrogen sulfide and carbon dioxide are present in the gas and the destruction occurs as a result of the chemical interaction between the pipe metal and these components, then we can talk about corrosion damage to the metal. The consequence of corrosion is always a change in the properties of the metal (chemical and mechanical), and as a result, a change in the operational characteristics of the pipeline. However, a pipe affected by corrosion may continue to function without failure for a certain period of time.

Cracks are a common type of defect in metal pipes. They can occur either longitudinally or annularly, and can be found in the base metal or at a welded joint. Our database takes into account all these possibilities. It should be noted that cracks are often caused by stress concentrators, such as non-welds, sinks, or other defects in the pipe or weld. Additionally, pressure drops in the pipe can also lead to cracks. Both of these factors can contribute to the formation of a crack in a pipe. The last type of defect in steel pipes that we will consider is erosion. This is a gradual loss of material from the pipe wall caused by the flow of liquid and particles that are transported through it. The impact of these particles on the wall and the resulting cavitation effects lead to local deterioration of the inner surface of the pipe.

A number of scientific studies, such as [17] and [18], have also conducted optical diagnostics of pipes using crawler robots, successfully detecting defects using machine learning techniques. Other studies obtained positive results by acquiring and processing magneto metric and acoustic data using neural networks [19]. The algorithms and neural networks applied demonstrated relatively high accuracy in detecting defects such as corrosion and cracks (about 69–75%), which is considered a good result for field studies, given the low quality of the initial data and limited statistical samples. Of course, the results obtained so far cannot be compared to those obtained by using machine learning on diagnostic samples in a laboratory setting. This is especially true for experiments on individual metal plates with defects. When learning occurs on such samples (cleaned of contaminants) under direct illumination, as well as on a flat, rather than a concave surface, the results of defect recognition can be extremely high.

For example, in [20], the proposed GAN-based augmentation scheme significantly improves the performance of CNN for classifying surface defects. The classically augmented CNN yields sensitivity and specificity of 90.28% and 98.06%, respectively. In contrast, the synthetically augmented CNN yields better results with sensitivity and specificity of 95.33% and 99.16%, respectively. Unfortunately, it is extremely difficult to obtain similar results in sensitivity and specificity in the field; this is probably not due to the quality of the proposed algorithms, but rather to a large number of external factors. Despite these factors, researchers are making quite successful attempts to increase the performance of a neural network and increase the accuracy of detecting various types of deviations, defects and objects [21]. However, the issue of testing the results obtained on a large volume of field data and the applicability of the developed algorithms and networks for solving specific diagnostic problems remain the subject of further research. A separate problem which is highlighted by a number of works such as [22] and [23], is the binding of a pipeline defect to a specific point along the route. It is important to not only determine the presence or absence of a defect, but also to indicate the problem area. In steel pipes, due to the possibility of using them as signal amplifiers, such a problem is not as severe [24]. But it represents a certain challenge for the detection and localization of defects in polymer reinforced and plastic pipes [25].

During the standard in-tube optical inspection, the flaw detector records all observed pipe defects, determining their geometric parameters and referring them to the picket along the route. In this case, binding was carried out to a specific frame that was received. At the same time, no selection of acceptable or unacceptable defects was made. All potential defects were recorded in the processed frame. Determining acceptable and unacceptable defects for a specific type of pipe requires creating an additional add-on in the software. To do this, it will first be necessary to determine the geometric parameters of the defect (information on frame size, focal length, illumination, etc., will need to be provided to the neural network) as well as information on acceptable defect sizes for this pipe type (a defect library linked to regulatory and technical documentation requirements must be created).

In general, this suggests the need for a three-stage development of an in-tube optical diagnostic system that uses neural networks to recognize defects [26]. At the first stage, the neural network would be trained to recognize defects from a live video feed (based on the results of this study). At the second stage, either a system would be developed to combat distortion (by creating a scan of the frame) or the neural network would be refined to recognize defects when shooting from side cameras (taking into account connection defects). Finally, at the third stage, a representative library would be assembled containing data on the geometric parameters of defects and their acceptability for pipes of different types and brands. This would allow for secondary processing of diagnostic results, not only to mark defective zones but also to determine the acceptability or unacceptability of a defect. Next, let's move on to solving the main goal of this research: recognizing defects in steel pipes from live video footage.

3. The Formation of a Neural Network and The Creation of a Library of Defects

As part of the study, a library of images containing a total of 500 images was used. These images of steel pipes were captured through in-line photography of different types of oil and gas steel pipes. This library served as the foundation for the development of a learning neural network. During the creation of this system, advanced techniques from the field of artificial intelligence, specifically machine learning, were employed.

So, the YOLO neural network was used as a basis. This network makes it possible to achieve high accuracy and speed in detecting defects in pipelines. The YOLO system, developed by Joseph Redmon and others, considers object detection as a regression task on spatially separated bounding boxes and associated class probabilities. It examines the entire image during testing, so its predictions are based on the global context of the image. The YOLO algorithm uses a convolutional neural network, which is a deep learning algorithm that takes an input image and assigns importance (studied weights and offsets) to aspects or objects in the image, distinguishing one from the other. Compared to other algorithms, images require much less preprocessing with this method.

Figure 3 shows the architecture of the YOLO network. The basic idea is to reduce the feature dimensionality using convolutional layers alternating with 1x1 convolutions that compress features from preceding layers. YOLOv8 consists of several key components. Backbone - this is the core part of the model designed to extract features from the input image. YOLOv8 uses a new Backbone architecture designed to improve the efficiency and accuracy of the model. Neck -this section connects the Backbone to the Head and is responsible for aggregating and passing on the extracted features. YOLOv8 uses structures such as Spatial Pyramid Pooling - Fast (SPPF) and Path Aggregation Network (PANet), which improve the model's ability to handle contextual information and different scales of objects. Head (output part) - this component is responsible for predicting the coordinates of bounding boxes, object classes, and assessing the confidence in detection. YOLOv8 uses a new Anchor-Free head, which simplifies the prediction process and improves the accuracy of the model.

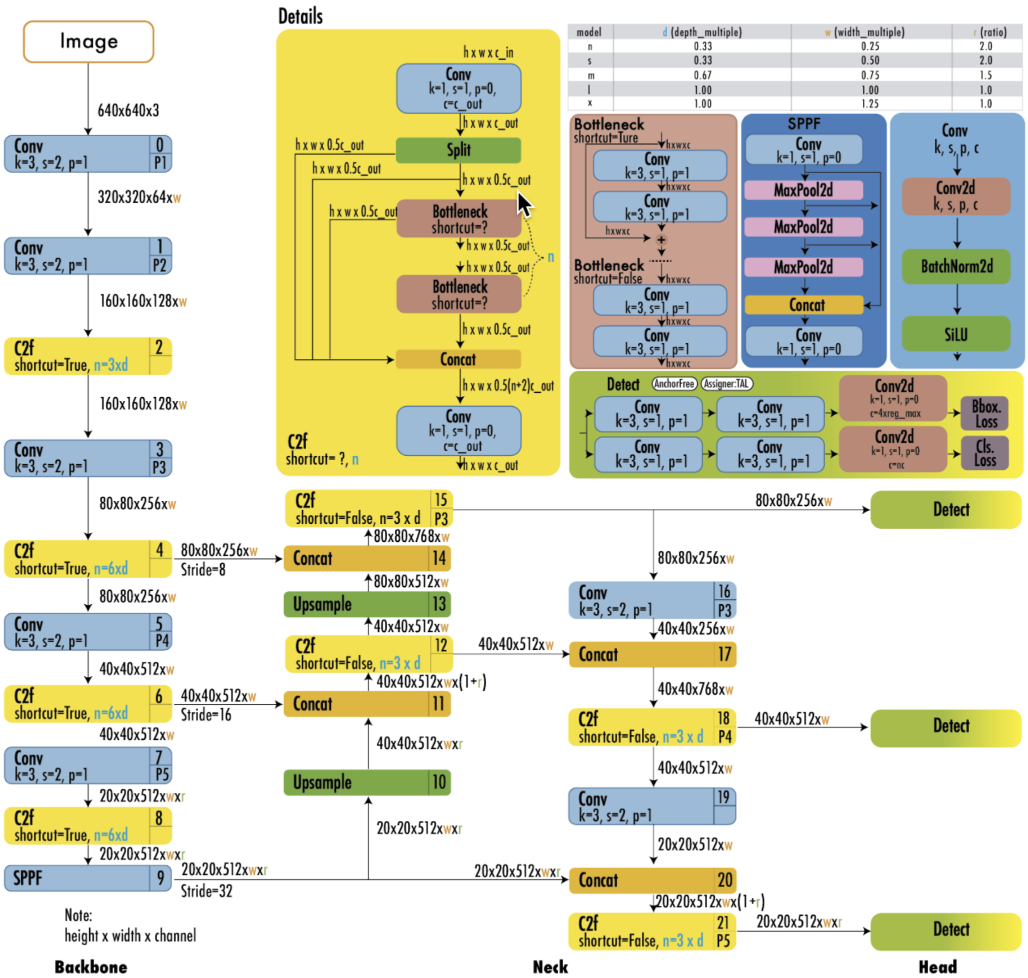

Fig. 3. Architecture of a convolutional neural network. The top right part of the diagram shows the model parameters: depth_multiple – model depth multiplier (0.33, 0.5, 0.75, 1.0, 1.25). Width_multiple – model width multiplier, ratio – scaling factors

Important improvements in YOLOv8 include the Anchor-Free approach. It involves abandoning the use of pre-defined anchor boxes, allowing the model to be more flexible and accurate in detecting objects of different sizes and shapes. A new loss function is also used, which contributes to more stable and efficient model training. An improved backbone network is also used. The updated Backbone improves the efficiency of feature extraction, which contributes to increased detection accuracy. The input image fed to the network has the size: 640×640×3 (RGB image with three channels). It is fed to convolutional layers (Conv), which reduce the dimensionality and extract features. The network processes this image through a series of convolutional layers, each of which performs a convolution using different kernel sizes. Convolutional layers reduce the dimensionality of the image and increase the number of channels. C2f (Cross Stage Partial with Fusion) is used - an improved version of Cross Stage Partial (CSP) blocks, responsible for efficient feature extraction and SPPF (Spatial Pyramid Pooling - Fast) – a layer that uses multiple levels of MaxPool2d to improve the model's perception of different scales of objects, allowing the model to focus on the most significant features [27].

YOLO convolutional layers are alternated with 1x1 convolutions, which help to reduce the number of features between main convolutional layers, thus reducing the computational load and increasing the efficiency of the model. MaxPooling layers with a 2x2 kernel and a stride of 2 are also used to reduce the image size, halving its width and height at each stage. This helps to preserve important features while reducing the image's size. The convolutional layers are followed by two fully connected layers. The first fully connected layer contains 4096 neurons, and the second contains 30 output neurons, which are responsible for predicting the coordinates of bounding boxes and object classes. These layers complete the image processing and form the final predictions.

Three outputs (P3, P4, P5) are responsible for detecting objects at different scales: P3 (80×80) – for small objects, P4 (40×40) – for medium objects, P5 (20×20) – for large objects. The network uses Conv2d convolutional layers to predict Bounding Box, Class, Objectness Score.

During the training process, visualization was performed to monitor the effectiveness of the training. The learning rate is a very sensitive parameter. With a very high learning rate, the error curve will have an unacceptable shape. With a low learning rate, the error will decrease very slowly even after a large number of epochs. With a high learning rate, the error initially decreases quickly but then becomes stuck in a local minimum, preventing the network from achieving good accuracy. When the learning rate is appropriately chosen, the error smoothly decreases to a minimal value.

Thanks to YOLO, it is now possible to quickly and accurately detect the location and type of defects, which is essential for a swift response to potential security risks. However, a significant amount of labeled data was required to train the neural network. Initially, the process of labeling defects in images was done manually. Later on, the Label Studio tool was utilized. With its assistance, it became feasible to swiftly create annotations for pictures of faulty pipes, which served as the foundation for training the neural network.

One of the key benefits of using YOLOv8 is its built-in optimized architecture, which greatly simplifies the process of training and deploying a model. Unlike alternative frameworks such as Detectron2 or MMDetection, which require explicit definition of neural network layers and training parameters, YOLOv8 offers a ready-made, well-tested architecture that provides a balance between accuracy and speed. The Ultralytics library provides a convenient API that allows us to run training, validation, and inference in just a few lines of code, reducing the complexity of development and making it accessible even to users with minimal experience in machine learning. In addition, YOLOv8 includes automatic hyperparameter tuning, built-in data augmentation functions, and computational optimization, making it an ideal choice for object detection tasks (Figure 4).

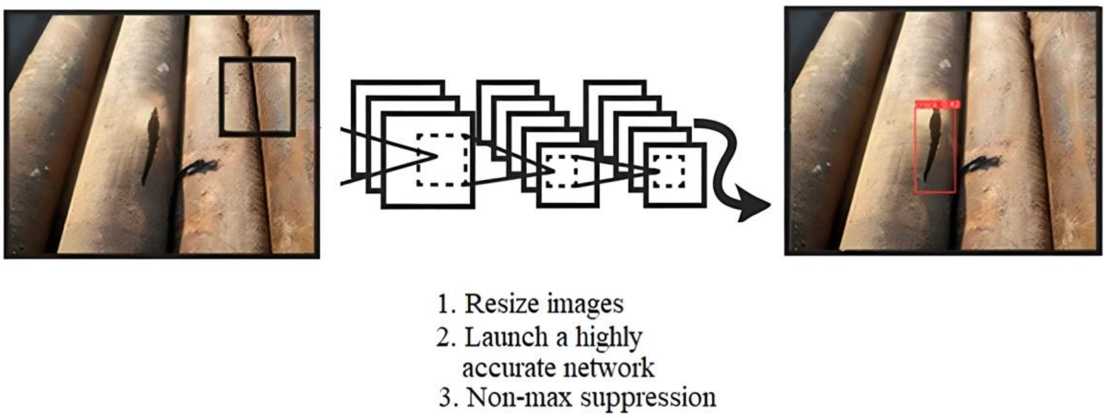

Fig. 4. YOLO object detection and convolutional neural network

Another advantage of YOLOv8 is its high speed, achieved thanks to its single-stage architecture, which is critical when analyzing large amounts of data in real time. Using pre-trained weights and built-in methods for fine-tuning allows the model to be adapted to various datasets with minimal computational resources. These features make YOLOv8 the most effective solution for object detection, especially in conditions where ease of integration, performance, and detection quality are important.

In this work, "YOLOv8" was used – the latest version of this model. Data preparation for training the YOLOv8n object detection model includes several stages. The first stage of data preparation is to collect and organize a set of images on which various objects are marked. The data is organized in a hierarchical folder structure, where each subfolder corresponds to a specific class of objects. All images are stored in a folder named "images". In the second stage, annotation tools such as "Label studio" are used to create annotation files in YOLO format. These files, usually with the extension ".txt", are saved in the "labels" folder and contain information about the coordinates of the defect boundaries (x, y, width, height) on each image. The third stage involves dividing the dataset into three parts: training (train), validation (valid) and test (test). The ratio between these parts is usually 70% for the training set, 15% for the validation set and 15% for the test set [28]. The structure of the prepared dataset is shown in Figure 5.

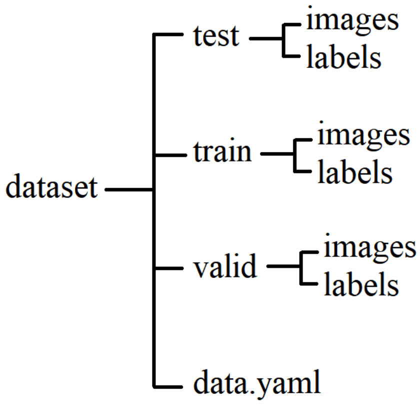

Fig. 5. Folder structure

At the fourth stage, the configuration file is created. The YAML configuration file plays a key role in configuring and managing the parameters of the YOLOv8n model. This file, usually with the extension ".yaml", contains detailed information about the structure of the model, as well as about the training parameters and data. Next, the model object is created. This file contains the weights of a model that has been pre-trained on a large dataset such as COCO, which allows you to use this model as a starting point for further training. The final step is to train the model.

4. Training a Neural Network to Detect Pipe Defects and Analyzing the Results

The training of the YOLO neural network was performed using the ultralytics library. To begin, the YOLO class was imported from the library using the code 'from ultralytics import YOLO'. This allowed us to utilize YOLO's features in our research. Then, an instance of the YOLO class was created with the pre-trained YOLOv8x model stored in the file 'yolov8x.pt'. The line of code used was 'model = YOLO("yolov8x.pt")'. After this, we proceeded to train the model (see Figure 6).

Starting the process of training the model using the 'train' method, we specified the path to the training data file, the number of training epochs, the image size, and the batch size. For example, the line of code: results = model.train('external_defects/crack/data.yaml', 250, 640, 16) starts training the model on the data from the 'external_defects/crack/data.yaml' file for 250 epochs, with an image size of 640 pixels and a batch size of 16. The results of the training are stored in the 'results' variable. This allows us to analyze and evaluate the performance of the model. Thus, based on the generated training and learning algorithms, all 500 images were processed. At the same time, groups of human-recognized images were previously formed with reference to defects of a specific type. Let's further explore the process of training a neural network and see its results.

During each iteration of training, the YOLO model processes images from the training dataset using a neural network. The images are analyzed and classified by the network, and features are extracted to determine the presence of objects. After analyzing the images, the model compares its predictions with actual image labels to calculate losses. These losses show how different the predictions are from the true values. The error is then propagated backwards through the network, where the weights and parameters are adjusted to minimize the error and improve predictions. This process is repeated multiple times during each epoch of training. An epoch ends when all weights in the network have been updated based on data from the training set. It is important to note that after each training iteration, the model's performance is evaluated on a separate validation dataset in order to monitor the learning process and prevent overfitting. Training stops when a certain stopping criterion is met, such as reaching a certain level of accuracy or completing a predetermined number of iterations.

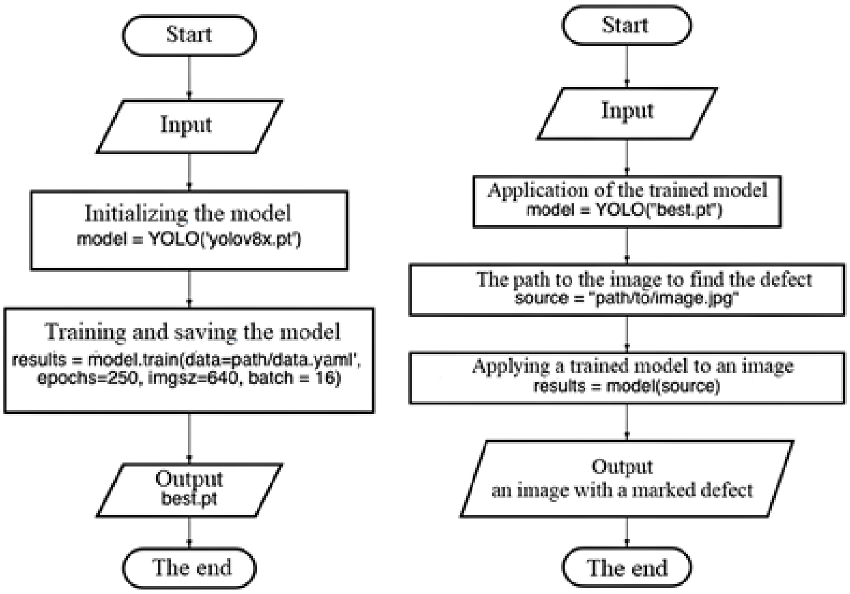

Fig. 6. Neural network training and learning algorithms

The model is trained using Python and the Ultralytics YOLOv8 library. The dataset is pre-labeled in the YOLO format, which includes object annotations in the form of bounding box coordinates and corresponding classes. We use data augmentation strategies to improve the model's robustness to various shooting conditions. The training code is provided in Appendix 2. During the training, hyperparameters such as the number of epochs, input image size, and mini-batch size were varied to optimize the detection quality. Using a GPU allowed us to speed up the training process and improve the convergence rate of the model. After training, the model is evaluated on a test dataset, where its accuracy (mAP) and defect detection completeness are analyzed. As a result of training the YOLOv8 model on a specialized dataset containing images of defects in steel oil and gas pipes, key metrics characterizing the detection accuracy and prediction quality were obtained. The training was carried out for 250 epochs, and the final learning rate was 0.00003584, which indicates gradual convergence of the model. The mean square error (MSE) on the training set was 1.7973, and on the validation set - 2.3834. Such values indicate the presence of a certain gap between the training and validation samples, which may indicate slight overfitting of the model. The regression error, coinciding with the MSE, confirms the stability of the model in predicting the coordinates of the bounding boxes of objects. The prediction accuracy was assessed using the mAP (Mean Average Precision) metric. The average mAP50 value, which measures the accuracy of object detection with a 50% intersection over Union (IoU) tolerance, was 0.2076, and the more stringent mAP50-95 metric, which takes into account different IoU thresholds, was 0.0988. This indicates that the model has potential for improvement, especially in terms of defect localization accuracy. The total error of the model, calculated as 1 - mAP50-95, was 0.9012, which indicates the need for further optimization of the parameters or an increase in the amount of training data. In addition, the precision and recall metrics were calculated. The average precision value was 0.3325, which means that 33.25% of the predicted objects actually correspond to real defects. However, the recall is 0.3036, which indicates that the model detects only about 30.36% of all defects present in the data. Such a gap between precision and recall indicates a tendency for the model to miss defects, which may be due to insufficient information content of the dataset or the complex structure of the detected objects.

Speaking about the results obtained, we will consistently focus on the various types of defects detected on the inner surface of the steel pipe. For example, the detection of bulge-type defects in pipes turned out to be insufficiently accurate. This was due, firstly, to the small number of available photographs containing defects of this type, which limited the number of examples on which the model could learn. Secondly, the quality of these photos turned out to be insufficient and this, in turn, could make it difficult for the neural network to recognize defects. Factors such as poor lighting, noise, distortion or blurring could have affected this. To improve the results of the study, it is necessary to pay attention to the process of taking images and take measures to improve the quality and quantity of photographs with a bug type defect (Figure 7).

Fig. 7. An example of training a neural network to recognize a bulge-type defects

The overall performance of YOLOv8 in the proposed data set for the bulge defect is illustrated by graphs of accuracy, average accuracy, response and errors for the models in Figure 8. As can be seen from the graphs, the model reaches no more than 48% accuracy, and up to 85% response. According to mAP, very low indicators are obtained, which indicates a weak detection of this type of defects during training.

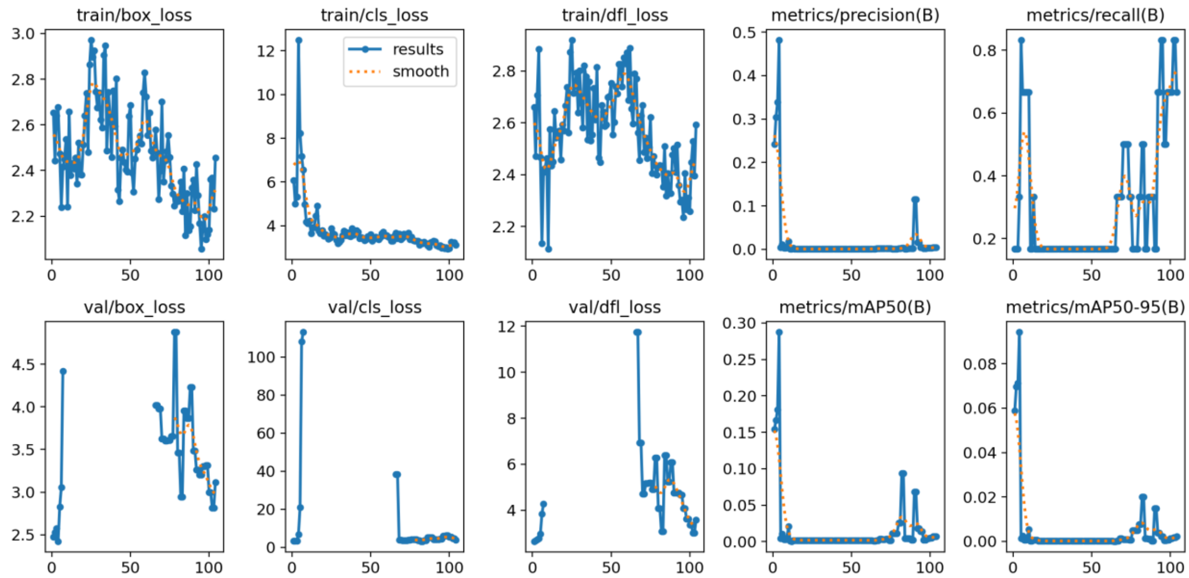

Fig. 8. The loss function and metrics of the YOLOv8 model in determining Bulge-type defects in steel pipes. The first line shows the learning curves for loss of regression in the block, loss of class, accuracy, and response. The second line shows the validation curves for block regression loss, class loss, average accuracy: 50 and average accuracy: 50:95

The results of training a neural network to detect a corrosion defect turned out to be significantly better than for a bulge defect. In general, a neural network can perform rather effective segmentation of defects of this type inside steel pipes. At the same time, the network made it possible to determine both areal corrosion, as well as pitting and stream corrosion. The learning outcomes are shown in Figure 9. In terms of training efficiency, it can be noted that sorting losses decreased from about 2.5 to 1.0. Classification losses showed good convergence from 5.0 to 1.0. Distribution Focal Loss (DFL) decreased from 2.5 to about 1.5. Accuracy reaches about 0.6, and recall time stabilizes at about 0.4. In terms of validation metrics, Val/box_loss stabilizes at 2.6-2.7, which is higher than in the erosion model, and Val/cls_loss drops to almost zero after early spikes. The final mAP50 is about 0.45, and mAP50-95 is about 0.15.

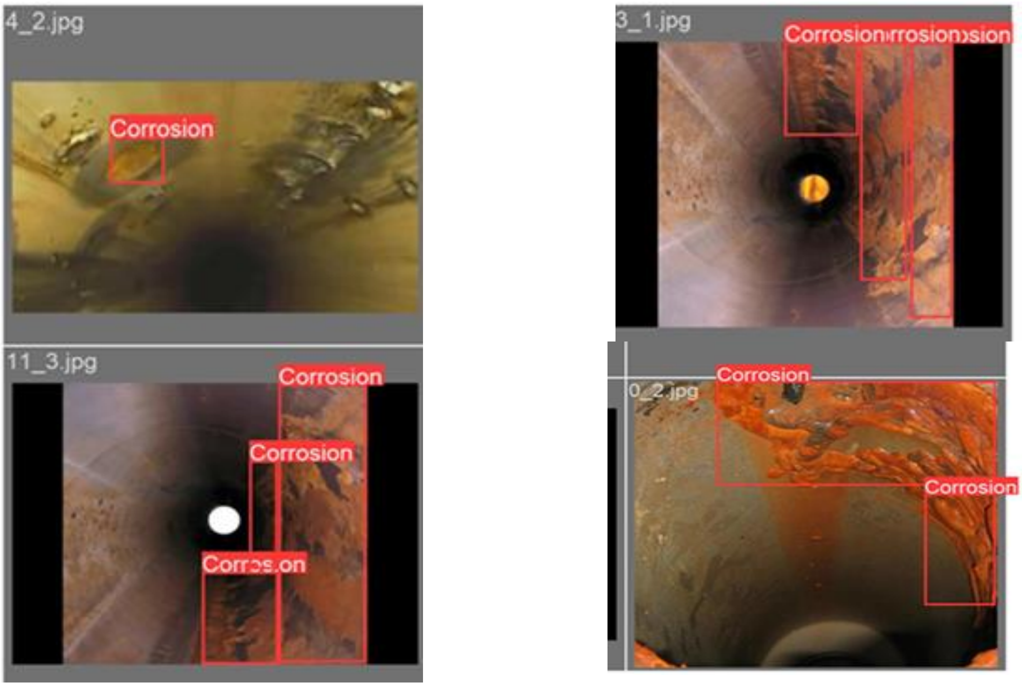

Fig. 9. An example of training a neural network to recognize a bulge-type defects

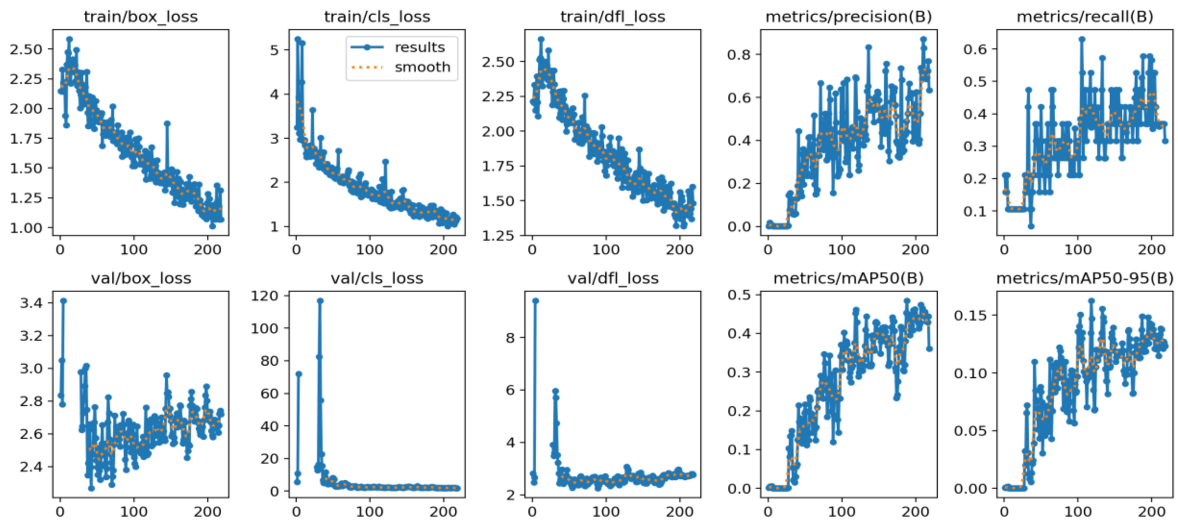

Fig. 10. The loss function and metrics of the YOLOv8 model in determining corrosion defects in steel pipes. The first line shows the learning curves for loss of regression in the block, loss of class, accuracy, and response. The second line shows the validation curves for block regression loss, class loss, average accuracy: 50 and average accuracy: 50:95

Overall, the results show a relatively high selectivity for the images. The corrosion defect is segmented much more accurately than the bulge defect. Compared to the accuracy of recognition and the results of the learning process for the bulge defect, the corrosion defect is segmented more accurately. This can be seen from the clearer trends in the charts with a smaller spread of individual values. Similarly, the performance indicators for YOLOv8 on the corrosion defect are summarized in a series of graphs (Figure 10). In this case, the accuracy of the model is higher, reaching 88%, while the response is lower compared to the previous type of defect but still reaches 67%. Overall, this gives a good result, with an average precision of 49% according to the mAP metric. For a crack-type defect, very good image segmentation results were also obtained (Figure 11). At the same time, the neural network made it possible to identify both longitudinal and transverse cracks of various types, and to isolate a grid of cracks. With an increase in the size of the Dataset, segmentation accuracy is expected to increase. Summary data of the loss function and metrics for the YOLOv8 model on the crack defect is shown in Figure 12. The model achieves an accuracy of up to 98% with a response rate of up to 57%, but the heterogeneity in prediction and response results in a relatively low mean average precision (mAP) of around 42%. This may be due to the need for more training data. It should be noted that for the full implementation of the results in real conditions, it is necessary to unify the illumination of the survey areas, expand the database and increase the amount of data for training the neural network.

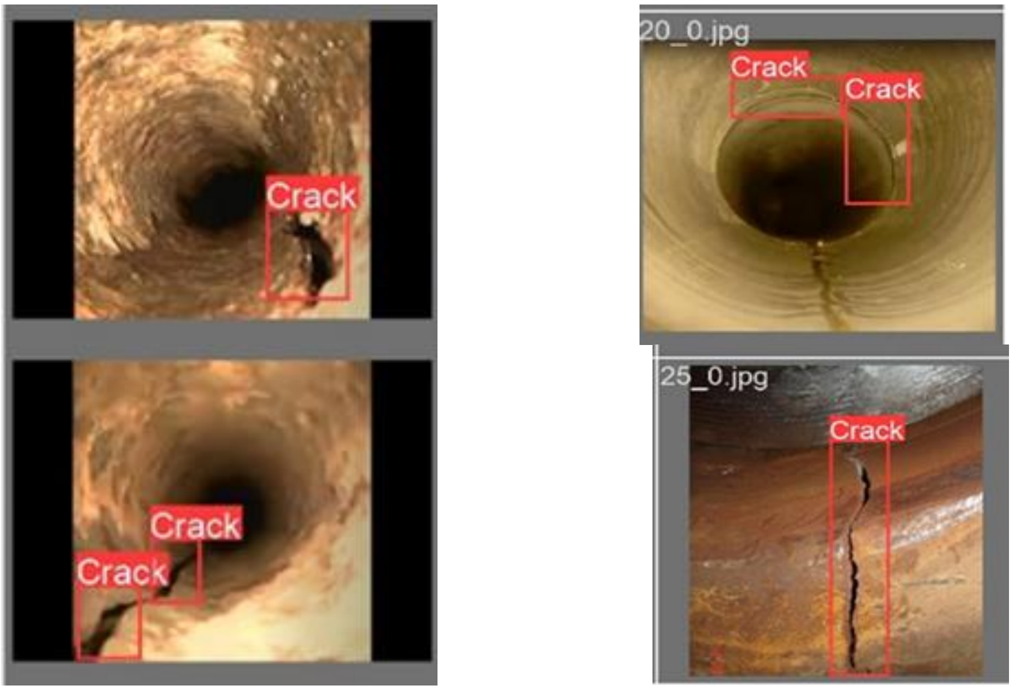

Fig. 11. Example of training a neural network to recognize a crack-type defect

For crack segmentation, the box loss (train/box_loss) shows a steady decrease from about 2.5 to 1.0 over 250 epochs, indicating good bounding box prediction convergence. The classification loss (train/cls_loss) has significantly decreased from about 5.0 to less than 1.0. The DFL loss (train/dfl_loss) shows a similar decreasing trend from 2.5 to 1.2. The precision and recall metrics show an increasing trend, with precision reaching about 0.5 and recall about 0.4 by the end of training. Analyzing the validation metrics, we note that for the crack dataset, the validation loss shows an initial spike followed by stabilization. The Val/cls_loss and val/dfl_loss metrics decrease sharply at the beginning of training. The mAP50 metrics steadily increase to about 0.35, and the mAP50-95 metrics gradually improve to about 0.2. A neural network to recognize an erosion type defect is presented (Figure 13). The results of the erosion defect analysis are summarized in a series of graphs (Figure 14). In this case, the model's accuracy is high, reaching 87%. This is an improvement compared to the previous type of defect, with a response of 68%. This makes erosion the most accurately detected defect, based on the mAP metric. The value reached 65%, which is the best result among all the considered defects. Looking at the training metrics for erosion, we see that the loss has decreased from 2.5 to about 0.5, which shows better convergence than the crack model. The classification loss has decreased from 5.0 to about 0.5, and the accuracy has reached higher values (about 0.7) compared to the crack model. The recall has improved to about 0.5. The validation metrics show more stable loss curves after the initial jumps. Importantly, the final val/box_loss has a lower value (about 2.4) compared to the crack model. The Val/cls_loss value stabilizes at a level close to zero. We also note the higher mAP50 (about 0.5) and mAP50-95 (0.25) values compared to the crack model.

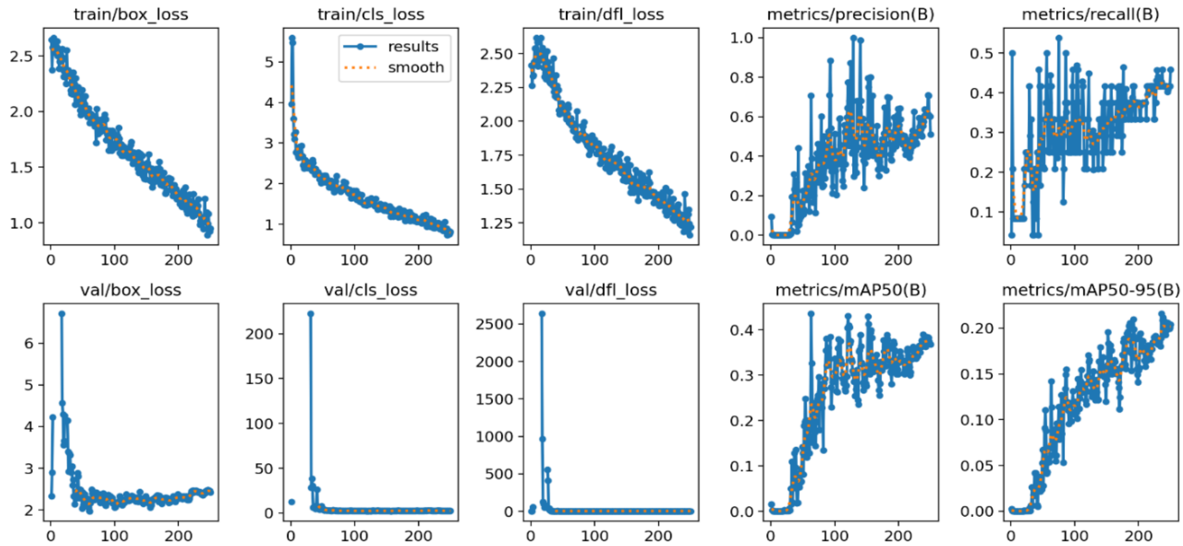

Fig. 12. The loss function and metrics of the YOLOv8 model in determining crack-type defects in steel pipes. The first line shows the learning curves for loss of regression in the block, loss of class, accuracy, and response. The second line shows the validation curves for block regression loss, class loss, average accuracy: 50 and average accuracy: 50:95

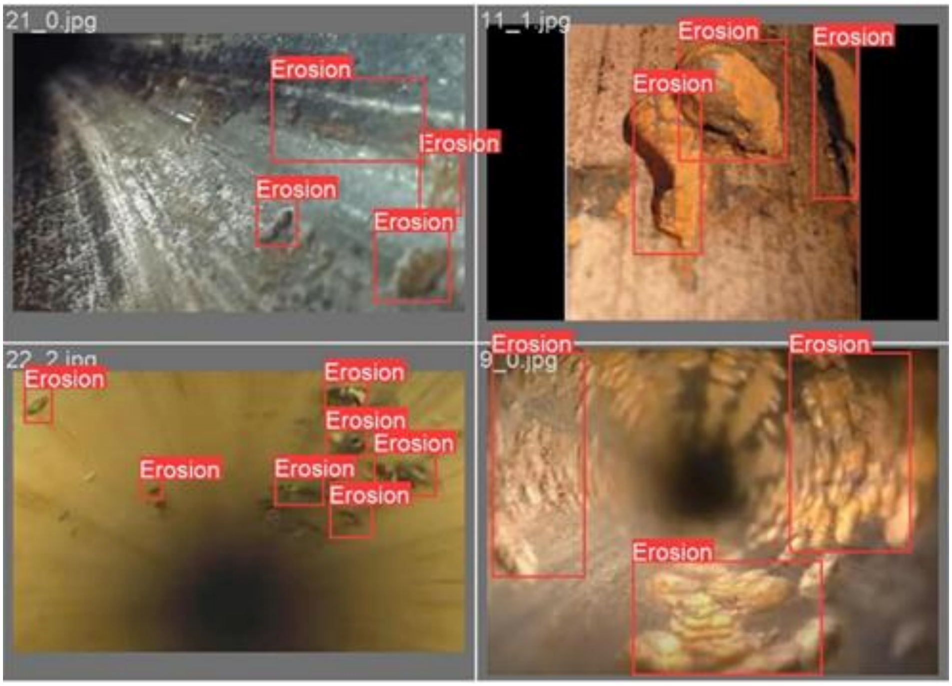

Fig. 13. Training a neural network to recognize an erosion type defect

The results obtained (Figures 8, 10, 12, and 14) indicate that, overall, neural networks are able to solve the tasks they are assigned. However, it is necessary to improve their accuracy and selectivity. It should be noted that the limited number of training data available negatively affects the model's performance [29]. Insufficient training examples make the model less capable of generalizing, which can result in low accuracy in defect recognition. Additionally, the quality of training data is also crucial [30]. Low-resolution or poorly lit photos of defects can make it difficult for the model to recognize them correctly. Based on the training results, we note that the erosion model demonstrates the best overall performance with the lowest final loss and the highest precision/recall and mapping scores. The crack detection model appears to be the most challenging due to its relatively high final loss values and low mAP scores. The corrosion model occupies an intermediate position between the other two models.

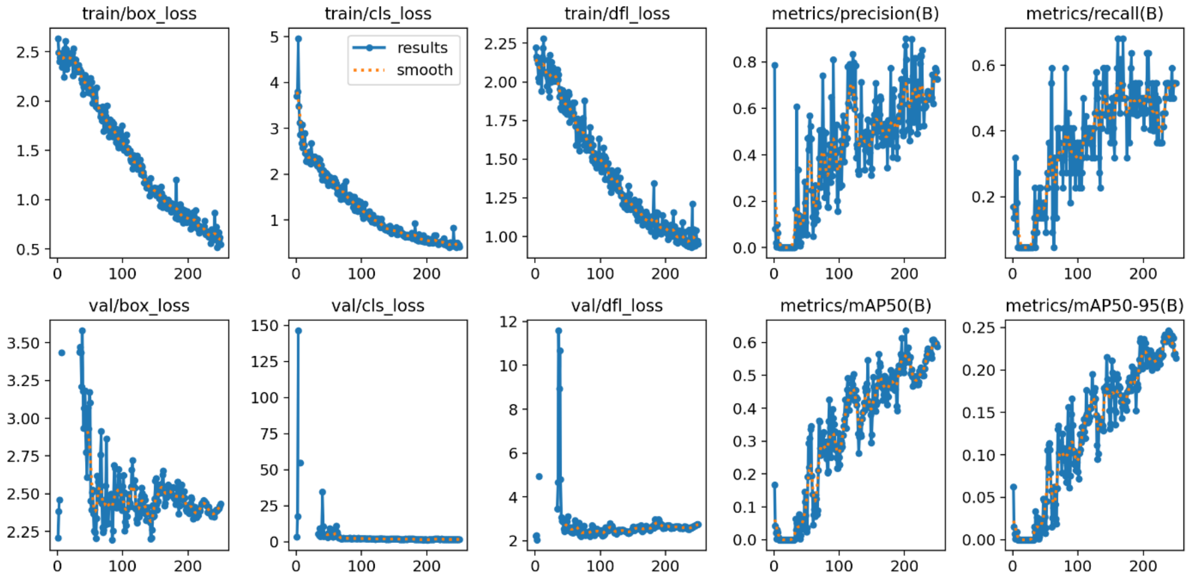

Fig. 14. The loss function and metrics of the YOLOv8 model in determining erosion-type defects in steel pipes. The first line shows the learning curves for loss of regression in the block, loss of class, accuracy, and response. The second line shows the validation curves for block regression loss, class loss, average accuracy: 50 and average accuracy: 50:95

Table 1. Results of training the neural network to recognize in-pipe defects in steel oil and gas pipelines that are under operation.

Type of defect	Precision	Recall	F1-metrics	mAP50	mAP50-95	Box Loss (train)	Box Loss (val)	Cls Loss (train)	Cls Loss (val)
Bulge	~0.0174	~0.30	0.032	~0.0068	~0.002	~2.46	~3.11	~3.1	~3.9
Crack	~0.50	~0.40	0.444	~0.35	~0.20	~1.0	~2.4	~0.8	~3.0
Erosion	~0.70	~0.50	0.583	~0.50	~0.25	~0.6	~2.4	~0.5	~2.0
Corrosion	~0.60	~0.40	0.480	~0.45	~0.15	~1.0	~2.7	~1.0	~2.5

All models display the expected training pattern: an initial high loss that decreases over time and evaluation scores that improve as training progresses. Performance metrics for the YOLO models include intersection over union (IoU), mAP, precision and recall, and F1 score. Summarize the results obtained for all types of defects in a single table (Table 1). As a result, a neural network was created that allows detecting and segmenting in-tube defects of the body surface in steel oil and gas pipelines that are under operation in the optical range. Based on this development, it will be possible to analyze both endoscopic examination data of steel pipes and probe-based examination data. Various defects are determined with different accuracy (the worst for buckles, the best for corrosion and erosion).

It is important to note that most of the results of neural network training for defect recognition presented in scientific periodicals are based not on data obtained from real operating conditions of technical systems, but on pre-prepared samples in a laboratory. Of course, the results, given in Table 1, indicate a relatively low selectivity of defects by the neural network. At the same time, they are based not on artificial and prepared laboratory pipe samples, but on photo and video materials obtained in operated field oil and gas pipelines. Contamination of pipe walls, difficult shooting conditions affected the quality of the dataset, which significantly increased the complexity of the images. With an increase in the sample size, and re-training of the neural network on a wider base, as well as using other types of neural networks in the future, it may be possible to obtain higher accuracy and precision in defect detection.

5. Conclusions

The conducted research establishes both the methodological framework and software basis for automating the in-line optical diagnostic process, with the potential to significantly reduce the cost of defect detection in steel pipes used in the oil and gas industry. The proposed solutions are applicable to optical inspection systems employing endoscopic equipment and robotic crawlers. The core of the approach is the application of convolutional neural networks (CNNs), specifically the YOLOv8 architecture, selected for its high detection speed and accuracy. From an applied perspective, the developed system has some selectivity in detecting defects on the inner surface of steel pipes in the optical spectrum.

The detection of the Bulge defect yielded the poorest performance among all defect types evaluated. Based on the current stage of research, it can be concluded that the neural network model is not yet capable of reliably identifying this defect in in-pipe optical diagnostics of operational oil and gas pipelines. These results indicate that the model fails to both accurately distinguish bulges from other image features and to localize them when detected. The model demonstrates moderate performance in crack recognition, with a precision of ~0.50 and recall of ~0.40, indicating a tendency toward false positives and a limited capacity to capture all relevant defects. The F1 score (~0.444) reflects a fair trade-off between precision and recall, suggesting foundational learning but insufficient reliability for deployment in safety-critical environments. Localization performance, as measured by mAP@50 (~0.35) and mAP@50–95 (~0.20), highlights challenges in accurately bounding crack regions, particularly under stricter IoU thresholds.

Moderately good results were obtained in erosion defect detection, with notable strengths in classification accuracy during training but clear limitations in generalization and precise localization. A precision of approximately 0.70 indicates high specificity, suggesting that the model effectively minimizes false positives. However, the recall of ~0.50 reveals a substantial proportion of missed erosion instances, limiting its reliability in scenarios where defect omission carries safety implications. The F1 score (~0.583) reflects a balanced, though not optimal, trade-off between precision and recall. Localization performance, as shown by mAP@50 (~0.50) and mAP@50–95 (~0.25), indicates acceptable detection under relaxed conditions but a drop in accuracy when stricter spatial alignment is required.

In case of corrosion recognition, the model exhibits partial effectiveness. The precision score of 0.60, indicates a moderate ability to correctly identify corrosion instances while maintaining a low false positive rate. However, the recall remains limited at ~0.40, reflecting the model's tendency to miss a significant proportion of true corrosion cases. The F1 score of ~0.48 highlights an imbalance between sensitivity and specificity. Localization performance under relaxed conditions is moderate (mAP@50 ≈ 0.45), but it declines sharply under stricter criteria (mAP@50–95 ≈ 0.15). For crack, erosion and corrosion defects detection, the difference between training and validation losses suggests possible overfitting. This point to the dataset being imbalanced. Field conditions and the limited number of images (at least 100 images per defect type) may cause such issues. We expect to obtain better results with the model as our dataset grows.

The lower detection rates observed for certain types of defects can be attributed to their complex geometries, uneven illumination, and the absence of side-view imaging, and the limited amount of images. The primary limitations affecting the deployment of the developed methodological and software framework include the small physical size of the target defects, the sensitivity of detection performance to specific lighting configurations, the necessity of pre-inspection pipe cleaning to prepare the surface for analysis, and the current incompleteness of the defect dataset. The latter underscores the need to expand the database to encompass a broader range of defect types commonly encountered in oil and gas pipelines. From a methodological perspective, the findings of this study have led to the identification of several promising avenues for further development, which could significantly enhance the performance and applicability of the proposed neural network model for defect recognition in steel pipelines. To improve the accuracy and robustness of defect detection, future work should focus on expanding the image dataset and add the ability to recognize new defect types.

A key future task involves the development of a unified and scalable database of permissible defect thresholds tailored to various steel pipe diameters and wall thicknesses. This database should be integrated with software capable of recalculating defect dimensions based on imaging parameters and standardized pipe geometries. Importantly, this system must align with current regulatory and technical standards governing the production and operation of oil and gas pipelines. Although the current neural network model demonstrates relatively modest accuracy, its potential remains high due to its training on real-world data obtained from operational oil and gas systems—data that is inherently unstructured and noisy, and thus more representative of field conditions than data obtained in controlled laboratory settings. This alignment with practical operating environments gives the model significant value and relevance for industrial application. To further advance this work, additional research should explore the incorporation of alternative video acquisition techniques, such as multi-angle or 3D imaging, which could improve detection coverage and spatial resolution. Finally, future investigations should aim to extend defect recognition capabilities beyond the pipe surface to include welding joints, which are critical areas of structural vulnerability.

Appendix 1. Table of abbreviations

YOLO	You Only Look Once
ConvNet/CNN	Convolutional Neural Network
mAP	Mean Average Precision
C2f	Cross Stage Partial with Fusion
SPPF	Spatial Pyramid Pooling - Fast
PANet	Path Aggregation Network
train	training
valid	validation
yaml	file type
ultralytics	YOLO Libraries
ISO	International standard organization
GOST	Russian State Standard
OST	Report (Industrial) Russian Standard

Appendix 2. Training code

```
from ultralytics import YOLO
# Loading a pre-trained model YOLOv8
model = YOLO("yolov8n.pt")
# Training on a custom dataset
results = model.train(data="dataset.yaml", epochs=50, imgsz=640, batch=16, device="cuda")
# Saving the trained model
model.save("yolov8_defects.pt")
```

References

[1] Kazeminasab S, Sadeghi N, Janfaza V, Razavi M, Ziyadidegan S, Banks MK. Localization, mapping, navigation, and inspection methods in in-pipe robots: A review. IEEE Access. 2021;9:162035-58. https://doi.org/10.1109/ACCESS.2021.3130233

[2] Sharkawy A-N, Ma'arif A, Furizal, Sekhar R, Shah P. A comprehensive pattern recognition neural network for collision classification using force sensor signals. Robotics. 2023;12:124. https://doi.org/10.3390/robotics12050124

[3] Al-Sabaeei AM, et al. Prediction of oil and gas pipeline failures through machine learning approaches: A systematic review. Energy Rep. 2023;10:1313-38. https://doi.org/10.1016/j.egyr.2023.08.009

[4] Malashin I, et al. Deep learning approach for pitting corrosion detection in gas pipelines. Sensors. 2024;24(11):3563. https://doi.org/10.3390/s24113563

[5] Hussain M, et al. Review of prediction of stress corrosion cracking in gas pi pelines using machine learning. Machines. 2024;12(1):42. https://doi.org/10.3390/machines12010042
[6] Chen J, Cao L, Song G. Detection of the pipeline elbow erosion by percussion and deep learning. Mech Syst Signal Process. 2023;200:110546. https://doi.org/10.1016/j.ymssp.2023.110546
[7] Liu L. Machine learning-driven corrosion detection and classification in pipelines. [dissertation]. University of Wales Trinity Saint David; 2024.
[8] Yuan J, et al. A classification and quantitative assessment method for internal and external surface defects in pipelines based on ASTC-Net. Adv Eng Inform. 2024;61:102492. https://doi.org/10.1016/j.aei.2024.102492
[9] Chen K, et al. An automatic defect detection system for petrochemical pipelines based on cycle-GAN and YOLOv5. Sensors. 2022;22(20):7907. https://doi.org/10.3390/s22207907
[10] Kammoun M, Kammoun A, Abid M. Leak detection methods in water distribution networks: A comparative survey on artificial intelligence applications. J Pipeline Syst Eng Pract. 2022;13(3). https://doi.org/10.1061/(ASCE)PS.1949-1204.0000646
[11] Zhu YK, et al. A review of optical NDT technologies. Sensors. 2011;11(8):7773-7798. https://doi.org/10.3390/s110807773
[12] Seghier MEAB, Höche D, Zheludkevich M. Prediction of the internal corrosion rate for oil and gas pipeline: Implementation of ensemble learning techniques. J Nat Gas Sci Eng. 2022;99:104425. https://doi.org/10.1016/j.jngse.2022.104425
[13] Hussain M, et al. Application of machine learning approaches to prediction of corrosion defects in energy pipelines. Adv Corros Model. 2024;127-166. https://doi.org/10.1007/978-3-031-60358-7_7
[14] Lyu F, et al. Application research of ultrasonic-guided wave technology in pipeline corrosion defect detection: A review. Coatings. 2024;14(3):358. https://doi.org/10.3390/coatings14030358
[15] Silva Motta R. Artificial neural networks for the rapid detection of defects. 2024.
[16] Petushkov G., Sigov A. Analysis and selection of the structure of a multiprocessor computing system according to the performance criterion. Russian Technological Journal. 2024;12(6):20-25. https://doi.org/10.32362/2500-316X-2024-12-6-20-25
[17] Zholtayev D, et al. Smart pipe inspection robot with in-chassis motor actuation design and integrated AI-powered defect detection system. IEEE Access. 2024. https://doi.org/10.1109/ACCESS.2024.3450502
[18] Fu Q, et al. Pipeline defect detection system based on crawling robots. In: Third International Conference on Electronic Information Engineering, Big Data, and Computer Technology (EIBDCT 2024); 2024. SPIE. p. 988-994. https://doi.org/10.1117/12.3031224
[19] Jain S, et al. Synthetic data augmentation for surface defect detection and classification using deep learning. J Intell Manuf. 2022;1-14.
[20] Liu Y, Bao Y. Review on automated condition assessment of pipelines with machine learning. Adv Eng Inform. 2022;53:101687. https://doi.org/10.1016/j.aei.2022.101687
[21] Sharkawy A-N. The effect of increasing hidden layers on the performance of the deep neural network: Modelling, investigation, and evaluation. Res Eng Struct Mater. 2024. http://dx.doi.org/10.17515/resm2024.442st0909tn
[22] Sampath S, et al. An innovative approach towards defect detection and localization in gas pipelines using integrated in-line inspection methods. J Nat Gas Sci Eng. 2021;90:103933. https://doi.org/10.1016/j.jngse.2021.103933
[23] Parlak BO, Yavasoglu HA. A comprehensive analysis of in-line inspection tools and technologies for steel oil and gas pipelines. Sustainability. 2023;15(3):2783. https://doi.org/10.3390/su15032783
[24] Hansen P, Alismail H, Rander P, Browning B. Visual mapping for natural gas pipe inspection. Int J Rob Res. 2015;34(4-5):532-558. https://doi.org/10.1177/0278364914550133
[25] Guo R, Tao Z. A study of neural network for surface characteristics in-process optical measurement. Optik. 2013;124(17):2821-2824. https://doi.org/10.1016/j.ijleo.2012.08.063
[26] Rizzo P, et al. Defect classification in pipes by neural networks using multiple guided ultrasonic wave features extracted after wavelet processing. 2005;294-303. https://doi.org/10.1115/1.1990213
[27] Datta S, Sarkar S. A review on different pipeline fault detection methods. J Loss Prev Process Ind. 2016;41:97-106. https://doi.org/10.1016/j.jlp.2016.03.010
[28] Mittal N, Vaidya A, Kapoor S. Object detection and classification using Yolo. Int J Sci Res Eng Trends. 2019;5(2).
[29] Kamenskiy A., Akaeva T., Grebenshchikova D. Digital three-stage recursive-separable image processing filter with variable sizes of scanning multielement aperture. Russian Technological Journal. 2024;12(6):48-58. https://doi.org/10.32362/2500-316X-2024-12-6-48-58
[30] Shcherban P, Sokolov A, Abu Hamdi R. Study of failure statistics of cavitators in the fuel oil facilities through the application of regression and cluster analysis. Proc Eng Sci. 2023;5(1):39-48. https://doi.org/10.24874/PES05.01.004