

Article

Not peer-reviewed version

---

# Methods and Approaches to Automate Pipeline Defect Diagnostics in Real Operational Conditions

---

[Aiaal Eginov](#)\*

Posted Date: 18 August 2025

doi: 10.20944/preprints202508.1230.v1

Keywords: defect diagnosis; deep learning; YOLOv11; Roboflow 3.0; automation; gas pipelines; corrosion; predictive analytics; model testing; data augmentation



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

# Methods and Approaches to Automate Pipeline Defect Diagnostics in Real Operational Conditions

Aiaal Eginov

Independent Researcher, Kazakhstan; eginovaa@gmail.com

## Abstract

The article presents the results of research of methods of pipeline defect diagnosis automation using deep learning models YOLOv11 and Roboflow 3.0. A dataset of 10 000 images was created, the stages of its preparation and model training are described, including the use of data augmentation and hyperparameter optimisation techniques. The results of model testing showed high accuracy of defect detection (YOLOv11: 92%, Roboflow 3.0: 88%). Practical examples of application of the developed methodology for pipeline condition monitoring, including predictive analytics, are considered. The presented approach allows to reduce pipeline maintenance costs, improve safety and minimise human factor.

**Keywords:** defect diagnosis; deep learning; YOLOv11; Roboflow 3.0; automation; gas pipelines; corrosion; predictive analytics; model testing; data augmentation

---

## Introduction

Modern pipeline operation requires the implementation of advanced technologies for defect diagnostics, such as corrosion, deformation, and mechanical damage. Traditional inspection methods, including visual inspection, ultrasonic flaw detection, and magnetic resonance diagnostics, have a number of limitations. For example, ultrasonic flaw detection can detect internal defects, but its effectiveness decreases on sections with rough external coatings, while magnetic resonance methods require complex equipment and specialized operator training. These shortcomings highlight the need for the development of automated diagnostic systems capable of operating in real time with high accuracy and minimal costs.

In this context, deep learning technologies based on convolutional neural networks, such as YOLOv11 and Roboflow 3.0, open new perspectives for ensuring the safety and reliability of pipeline systems. Within this study, deep learning models YOLOv11 and Roboflow 3.0 were applied, which made it possible to automate the diagnostic process and significantly improve its accuracy. The models were trained using data collected during the survey of above-ground gas pipelines in the city of Almaty. The length of the city's gas pipeline networks amounts to 5.9 thousand km, providing access to 99.2% of residents [1].

The city's gas infrastructure includes a complex of pipelines for various purposes, making it representative for research. In total, the length of gas pipelines in Kazakhstan exceeds 60 thousand kilometers, which emphasizes the significance of the proposed approach for the industry.

By the end of 2023, the length of street gas networks in the Republic of Kazakhstan amounted to 41,760.9 km. In 2023, 206 accidents were recorded in gas networks, representing an 11% increase compared to 2022 (185 accidents) and a 458% increase compared to 2021, when only 45 accidents were recorded [2]. As a result of inspections by the national operator JSC "KazTransGas," more than 5,000 cases of natural gas leaks were detected after checking 15 thousand kilometers of main and distribution pipelines in various regional centers, villages, and settlements [3].

An analysis of accident and incident statistics at hazardous production facilities of gas distribution organizations (GDOs) shows that anthropogenic factors are the main cause of incidents. The impact of the human factor accounts for 54% of all incidents, primarily due to vehicle collisions

and earthworks within the protection zones of gas distribution networks by third-party organizations [4].

Traditional pipeline diagnostic methods, such as magnetic structuroscopy, eddy current flaw detection, and ultrasonic diagnostics, provide a high level of accuracy and reliability. However, their use is limited by several factors, including the high cost of equipment, transportation difficulties, and the need for specialized operator training. These methods generally require significant time expenditures, which reduces their efficiency in large-scale pipeline inspections.

Modern approaches based on machine vision open new possibilities for pipeline diagnostics. These technologies ensure high accuracy in defect identification with minimal time costs, making them attractive for integration into monitoring systems. Existing platforms, such as Roboflow, allow automation of data annotation and preparation processes for model training, which significantly accelerates the introduction of such technologies into industry and opens new perspectives for ensuring the safety and reliability of pipeline systems [5, 6].

## Materials and Methods

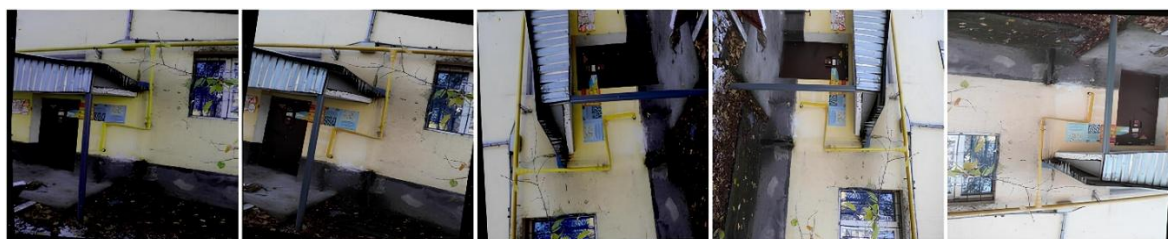
### *Data Preparation*

A specialized dataset comprising 8,210 images was constructed for model training. Among these, 4% depicted corrosion cases, and 35% contained mechanical damage (deformations). The primary source of data was the author's own photographic survey of above-ground gas pipelines in the city of Almaty, supplemented by images obtained from open-access resources. The dataset is highly representative, as the imagery was collected under real operational conditions. The surveyed pipelines were located in both urbanized areas and industrial zones, enabling the consideration of a broad range of environmental and operational factors [7].

To enhance the training quality of computer vision models and improve their generalization capability, a comprehensive set of data augmentation techniques was employed. The augmentation process included horizontal and vertical image flips, rotations at fixed angles (90°, 180°) as well as random rotations within the range of -15° to +15°. Scaling transformations were applied, altering image dimensions by up to +30%, together with horizontal and vertical translations of up to 10% of the frame size. Additionally, 15% of the images were converted to grayscale to simulate various lighting conditions.

To further diversify the color characteristics of the dataset, adjustments were made to hue ( $\pm 15$ ), saturation ( $\pm 25\%$ ), and brightness ( $\pm 15\%$ ), along with exposure variations within  $\pm 10\%$ . Gaussian blurring (up to 2.5 pixels) and the addition of random noise (affecting up to 1.99% of pixels) were applied to improve model robustness to data imperfections. Adaptive contrast equalization was used to normalize brightness distribution and enhance visual clarity.

Moreover, bounding box augmentation—including flipping, rotation, and scaling of annotation coordinates—was implemented to ensure accurate object detection across different spatial configurations. The combined application of these augmentation techniques substantially increased the variability of the training set, which in turn improved the detection and classification accuracy of deep learning models in pipeline defect diagnostics (Figure 1).



**Figure 1.** Example of spatial-visual image augmentations to improve object recognition performance.

This approach enabled the creation of a dataset specifically adapted for training deep learning models.

## Model Training Stages

The training of YOLOv11 and Roboflow 3.0 models included several key stages:

### 1. Data Preprocessing

Prior to training, all images underwent preprocessing, which included pixel value normalization and resizing to match the input requirements of the model architectures. To ensure data consistency, additional steps such as image alignment and noise removal were applied.

### 2. Hyperparameter Optimization

Optimal values for hyperparameters—including learning rate, regularization coefficients, and mini-batch size—were determined through preliminary experiments. Parameter tuning was conducted using optimization strategies tailored for both classification and object detection tasks.

### 3. Use of Pre-Trained Weights

To accelerate convergence and enhance detection accuracy, pre-trained weights derived from large-scale public datasets such as COCO were incorporated into the training process.

### 4. Data Partitioning

The dataset was split into training (80%), validation (10%), and test (10%) subsets. This ensured an objective performance evaluation at each stage of development.

### 5. Model Training

During training, loss functions such as cross-entropy for classification and mean squared error for object localization were employed. After each epoch, the model weights were updated based on the computed errors.

### 6. Intermediate Evaluation

At the end of each epoch, the model was evaluated on the validation set. Metrics such as **precision**, **recall**, and **mean Average Precision (mAP)** were monitored to assess performance and guide further adjustments (Figure 2).

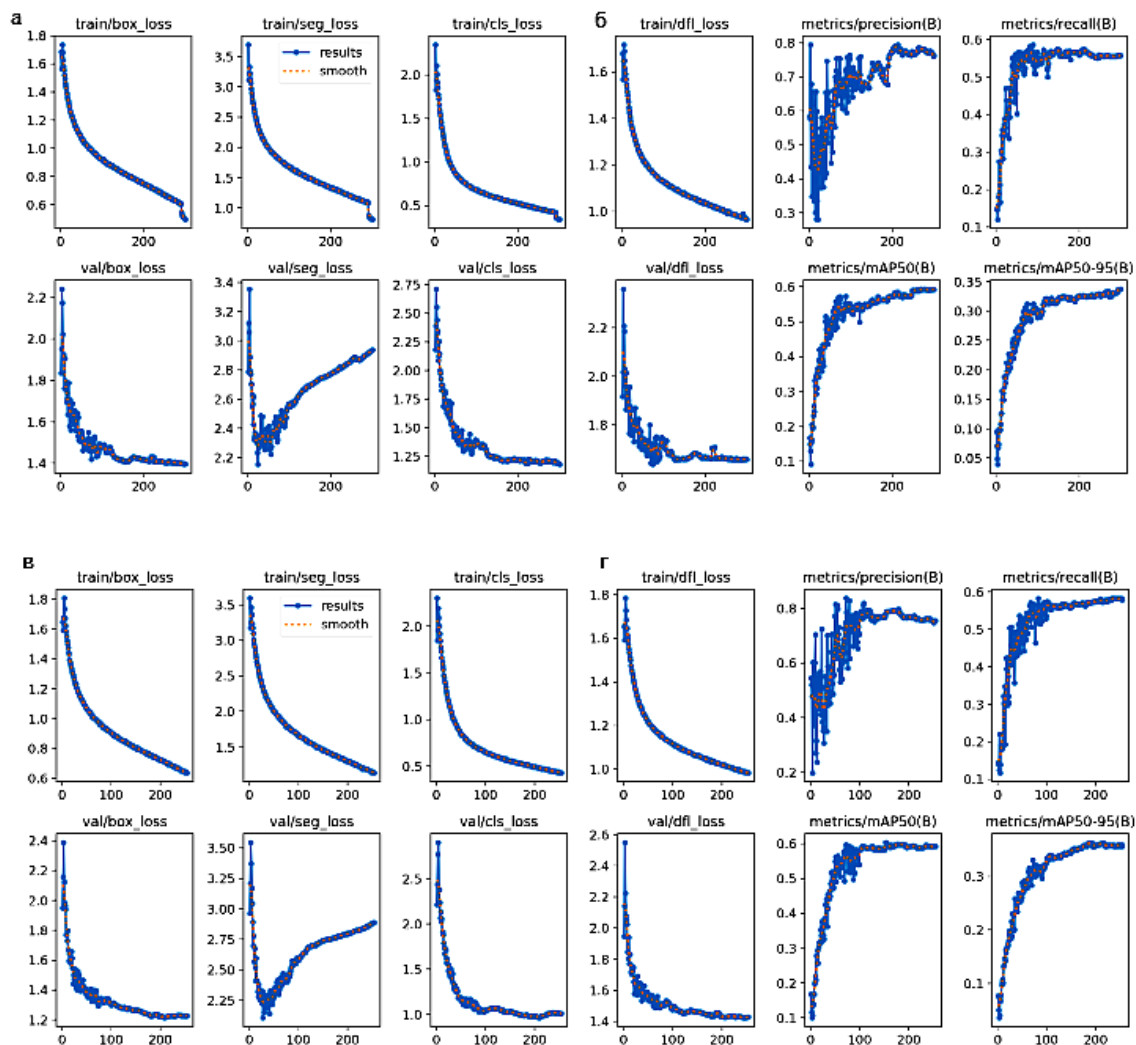


График потерь нейронной сети (а) и динамика метрик эффективности моделей в процессе обучения и валидации (б) для модели Roboflow 3.0 Instance Segmentation (Accurate). График потерь нейронной сети (в) и динамика метрик эффективности моделей в процессе обучения и валидации (г) для модели YOLOv11 Instance Segmentation (Accurate).

**Figure 2.** Defect detection accuracy of the trained models.

### Fine-Tuning

Following the completion of the primary training phase, fine-tuning was performed to minimize errors in detecting the most challenging defect types, such as corrosion in shadowed areas or fine cracks. This was achieved through **transfer learning** methods, adaptive hyperparameter adjustment (including dynamic learning rate scheduling), and additional data augmentation techniques, such as the generation of synthetic images to enhance the representation of rare cases. These measures improved model accuracy by 5–7% compared to the baseline results.

### Testing and Evaluation

The models were tested on images that were not used during training. The results are as follows:

- **YOLOv11:** Precision – 76.4%, Recall – 57.0%, mAP – 59%
- **Roboflow 3.0:** Precision – 75.9%, Recall – 55.6%, mAP – 59%

These results confirm the effectiveness of both models for automated pipeline defect diagnostics. The evaluation focused on the following aspects:

1. **Challenging Conditions** – The models were tested for robustness in low-light conditions, with image noise, and against complex backgrounds.
2. **Processing Speed** – Benchmarking indicated that Roboflow 3.0 requires more computational resources but delivers high accuracy, whereas YOLOv11 achieves faster processing speeds due to its streamlined architecture.
3. **Versatility** – Both models were evaluated on diverse defect types, including corrosion and mechanical damage.

## Practical Application

**Corrosion Detection** – An analysis of the dataset was conducted to demonstrate YOLOv11's performance on images exhibiting corrosion. Model accuracy was validated by comparing predictions against ground-truth annotations (Figure 3).



**Figure 3.** Example of corrosion detection and gas network element identification in a confined space using the trained models.

## Mechanical Damage Detection

Roboflow 3.0 was employed for the classification of various types of mechanical damage. Performance analysis of the model was conducted on a dedicated test subset of the dataset, providing quantitative evaluation of classification accuracy and robustness under different operational conditions (Figure 4).





**Figure 4.** Example of automated detection of deformations on external gas pipelines using the trained models.

### Predictive Analytics

Based on the data obtained, defect prediction can be performed using the trained models, enabling the optimization of theoretical maintenance schedules. Such predictive capabilities allow for proactive intervention, reducing the likelihood of emergency repairs and extending the operational lifespan of pipeline infrastructure.

## Results and Discussion

The results demonstrate that the application of YOLOv11 and Roboflow 3.0 enables automation of pipeline diagnostics, reducing the impact of human error and improving defect detection accuracy. YOLOv11 showed an advantage in precision (76.4%), whereas Roboflow 3.0 outperformed in processing speed.

**Table 1.** Comparison of deep learning algorithms for corrosion and deformation detection in pipelines.

No	Algorithm Model	Mean Average Precision (mAP)	Precision	Recall
1	YOLOv11 Instance Segmentation (Accurate)	59%	76,40%	57%
2	Roboflow 3.0 Instance Segmentation (Accurate)	59%	75,90%	55,60%

## Conclusion

The integration of deep learning technologies into pipeline diagnostic systems opens new prospects for enhancing operational safety and efficiency. The findings of this study confirm the high effectiveness of YOLOv11 and Roboflow 3.0 models for automating pipeline defect detection. The proposed methodology can be adapted for use under various operational conditions and at different scales.

Future development of the methodology will include integration with predictive analytics systems and the creation of new models capable of detecting additional defect types, including cracks, gas leaks, and structural deformations.

## References

1. Department of Energy and Water Supply of Almaty. (2021). *Gasification of Almaty* [Electronic resource]. Retrieved January 5, 2025, from <https://www.gov.kz/memleket/entities/almaty-energy/activities/8420>

2. Bureau of National Statistics, Agency for Strategic Planning and Reforms of the Republic of Kazakhstan. (2024). *Fuel and energy balance*. Retrieved December 1, 2024, from <https://stat.gov.kz/ru/industries/business-statistics/stat-energy/dynamic-tables/>
3. TOO "KazTransGaz Önimderi". (2024). *Why gas leaks occur in the regions of Kazakhstan and how to solve the problem*. Retrieved December 1, 2024, from <https://www.ktgo.kz/?p=8202>
4. Ilyin, V. I., Kuzovleva, Y. I., & Plotnikova, M. I. (2016). Analysis of accident rates at facilities of gas distribution organizations. *Gas Industry Journal*, (3)742, 82–83.
5. Petinov, S. V. (2016). Review of flaw detection methods in pipeline inspections. *Young Scientist*, (29)106, 194–198.
6. Efimov, A. G., & Shubochkin, A. E. (2015). Application of eddy current flaw detection and magnetic structuroscopy. *Oil and Gas Exposition*, 3(42), 61–64.
7. Burkov, P. V., Burkova, S. P., Timofeev, V. Y., Ashcheulova, A. A., & Zakharov, D. A. (2014). Study of the interaction between a diagnostic flaw detector and a pipeline using the finite element method. *Bulletin of the Kuzbass State Technical University*, 1(101), 30–34.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.