

Львівський національний університет імені Івана Франка  
Міністерство освіти і науки України

Кваліфікаційна наукова  
праця на правах рукопису

**Мисюк Роман Володимирович**

УДК 004.4:004.62

**ДИСЕРТАЦІЯ**

**МЕТОДИ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ЗМІНИ СТАНІВ ДЕФЕКТІВ  
НА ПОВЕРХНІ У ЕЛЕМЕНТАХ ІНЖЕНЕРНИХ КОНСТРУКЦІЙ**

122 — Комп'ютерні науки

12 — Інформаційні технології

Подається на здобуття ступеня доктора філософії

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело \_\_\_\_\_ (Р. В. Мисюк)

Науковий керівник Юзевич Володимир Миколайович,  
доктор фізико-математичних наук, професор

Львів – 2024

## АНОТАЦІЯ

*Мисюк Р.В.* Методи інтелектуального аналізу зміни станів дефектів на поверхні у елементах інженерних конструкцій — Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 122 «Комп'ютерні науки» (галузь знань 12 «Інформаційні технології»). — Львівський національний університет імені Івана Франка, Львів, 2024.

Своєчасна ідентифікація та прогнозування змін станів дефектів на поверхнях в інженерних об'єктах дозволяє запобігти руйнуванню конструкції та продовжити функціонування матеріалів. Під час аналізу пошкоджень у елементах конструкцій застосовуються різні підходи та апаратні системи, а також програмні засоби. Зазвичай відповідні комплексні системи забезпечують повний цикл опрацювання даних від моменту збору до їх візуалізації та аналізу. Для більш точного діагностування дефектів інженерних об'єктів використовуються різноманітні методики, включаючи ультразвукові, магнетоакустичні та вібраційні методи, а також обробку зображень, що базується на теплоконтролі, дефектоскопії та корозіометрії.

Постійне опрацювання великої кількості даних про поточний стан досліджуваного об'єкта (зокрема, параметрів) часто потребує збереження та вивчення попередніх історичних даних. Варто відзначити в цьому контексті, що з року в рік відбувається покращення ефективності мікропроцесорних пристроїв, зростання обсягів даних, збільшення кількості відповідних інформаційних технологій (ІТ), що дозволяє оптимізувати процес аналізу даних та здійснювати моніторинг змін параметрів системи.

Серед методів інтелектуального аналізу даних розглядаємо підхід із застосуванням алгоритмів комп'ютерного зору для більш точної сегментації

поверхневих дефектів різних інженерних об'єктів. Автоматизація процесу діагностики пошкоджень досліджуваних об'єктів відбувається з допомогою камери приєднаної до автономного пристрою. У випадку моніторингу елементів конструкцій типу підземних трубопроводів розглядаємо поверхневі дефекти (зокрема, корозійні та типу тріщин), для ідентифікації яких використовуємо відомі реалізації нейронних мереж. З метою оптимізації процесу та часу навчання моделей машинного навчання здійснювалося з використанням графічного процесора Nvidia GeForce RTX 3080 Ti у лабораторії технологій штучного інтелекту Львівського національного університету імені Івана Франка.

Апаратні засоби допомагають отримувати інформацію про вплив середовища на зміни поточного стану дефектів на поверхні елементів конструкцій (зокрема, трубопроводів) за допомогою зовнішніх сенсорів, з урахуванням комунікації між пристроями на основі бездротових технологій передачі даних (Wi-Fi, Bluetooth тощо). В цьому контексті доцільно використати методи та методики Технології Інтернету речей (IoT), оскільки вони дозволяють передавати множину зібраних даних із сенсорів (наприклад: інформацію про відносну вологість та кількість підрахованих пікселів у кадрі) до центрального сховища даних для подальшого їх аналізу.

Програмування діагностичних та моніторингових систем для аналізу пошкоджень (зокрема, дефектів) потребує формування низки статистичних моделей та методів для оцінювання критичних параметрів, які характеризують стан матеріалу елементів конструкцій. Методи машинного навчання дозволяють на основі гнучких моделей автоматизувати процеси класифікації та прогнозування щодо змін стану дефектів. Перевірку результатів прогнозу здійснюємо на основі вищезгаданих методик розпізнавання з допомогою мікрокомп'ютера Raspberry Pi та комплекту Nvidia JetRacer Kit.

З аналізу останніх досліджень із відкритих джерел випливає, що відомий обсяг результатів для аналізу та прогнозування зміни стану дефектів елементів

конструкцій на основі статистичних моделей за допомогою інформаційних технологій (ІТ) з урахуванням методів комп'ютерного зору недостатній для того, щоб можна було використати їх для опрацювання інформації про дефекти на поверхні інженерних конструкцій (зокрема, підземних трубопроводів). Тому варто застосовувати методики прогнозного аналізу для опрацювання зчитаних даних із сенсорів (камери, вологість, температурні зміни). Відповідний підхід дає змогу дослідити ефективність широкого спектру методів та методик розпізнавання зображень дефектів та аналізувати зміни в елементах інженерних конструкцій (ЕІК) з урахуванням поточних кліматичних умов.

Слід відзначити, що дисертаційна робота присвячена використанню методів та засобів інтелектуального аналізу змін станів поверхневих дефектів у елементах конструкцій (зокрема, у трубопроводах, інженерних об'єктах). В цьому контексті у дисертації також запропоновано варіанти застосування підходів фізико-математичного, статистичного та прогностичного моделювання для оцінювання критичного стану поверхневих дефектів для кіберфізичної системи “МПТ – СКЗ” (МПТ – металевий підземний трубопровід, СКЗ – система катодного захисту).

У першому розділі дисертаційної роботи розглянуто результати використання інформаційних технологій щодо розпізнавання дефектів у матеріалах з використанням робототехнічних засобів. Проаналізовано сучасний стан та аспекти розпізнавання дефектів на основі згорткових нейронних мереж та крайових фільтрів.

Визначено актуальні проблеми проектування ІТ з урахуванням концепцій використання методів комп'ютерного зору для аналізу зображень дефектів у темних місцях (тобто на внутрішній поверхні труб). Вагомий внесок щодо розвитку підходів, методів, моделей та інформаційних технологій на основі обробки зображень у контексті дефектоскопії та аналізу стану дефектів на поверхні елементів конструкцій полягає у впровадженні сучасних інструментів і

підходів для точнішої та ефективнішої ідентифікації дефектів та автоматизації процесу їх аналізу.

У другому розділі дисертаційної роботи проведено статистичний аналіз та здійснено прогнозування зміни стану дефектів. Для опису основного набору даних про аварійність та механічні пошкодження у процесі експлуатації трубопроводів використано звіти адміністрації з безпеки трубопроводів і поведінки небезпечних матеріалів за інформацією Міністерства транспорту США, а також аналогічні дані, отримані в Україні. Визначено основні залежності та встановлено кореляційні зв'язки стосовно параметрів впливу на аварійність елементів конструкцій, зокрема, таких як: зовнішні впливи, вік матеріалу, тривалість експлуатації та наявність ремонтів. Проведено порівняння параметричних та ймовірнісних методів для визначення змін стану дефектів і оцінювання ефективності та точності прогнозу. Використання отриманих результатів дозволяє більш комплексно діагностувати зміну стану елемента конструкції та оптимізувати систему антикорозійного захисту цього об'єкту.

Третій розділ дисертації присвячений апаратно-програмній реалізації прототипу ІТ для валідації використаних методів інтелектуального аналізу даних та оцінювання зміни станів дефектів у інженерних об'єктах.

Ідея розробки полягає в синтезі та дослідженні методів розпізнавання дефектів з використанням мікрокомп'ютера Nvidia Jetson Nano та аналізі дефектів на основі статистичного моделювання відповідних параметрів у контексті прогностичної моделі. Проведено огляд та аналіз методів бездротової передачі даних та опрацювання їх безпосередньо на мікрокомп'ютері для забезпечення автономності системи.

У четвертому розділі запропонований підхід вибору аналітичних засобів для фізико-математичного моделювання характеристик якості антикорозійного захисту конструкцій і застосування відповідних результатів для розробки ІТ. В основу відповідної інформаційної технології покладено фізико-математичну

модель поверхневих явищ та методи оцінювання якісних інформативних параметрів. В результаті реалізації ІТ розроблено новий удосконалений комплексний варіант оцінювання ресурсу для системи “металевий підземний трубопровід (МПТ) – СКЗ” у контексті комп’ютерного моделювання і враховує при цьому оптимізаційні критерії, методи та методика оцінювання змін параметрів та інформаційних потоків.

**Ключові слова:** інтелектуальний аналіз даних, машинне навчання, комп’ютерний зір, розпізнавання, Інтернет речей, обробка зображень, регресійні моделі, кіберфізична система, інформаційний потік, REST API, Raspberry PI, Nvidia Jetson Nano, тріщини, пітинги, дефектоскопія.

## ABSTRACT

*Mysiuk R.V.* Methods of intellectual analysis of changes in the state of defects on the surface of elements of engineering structures — Qualifying scientific work on manuscript rights. Dissertation for obtaining the scientific degree of Ph.D in specialty 122 "Computer Science" (12 "Information technologies"). — Ivan Franko National University of Lviv, Lviv, 2023.

Timely identification and forecasting of changes in the state of defects on the surfaces of engineering objects allows to prevent the destruction of the structure and to continue the functioning of the materials. Various approaches and hardware systems, as well as software tools, are used in the analysis of damage to structural elements. Usually, the corresponding complex systems provide a full cycle of data processing from the moment of collection to their visualization and analysis. Various methods are used for more accurate diagnosis of defects in engineering objects, including ultrasonic, magnetoacoustic and vibration methods, as well as image processing based on thermal control, defectoscopy and corrosiometry.

Continuous processing of a large amount of data on the current state of the object under study (in particular, parameters) often requires the preservation and study of previous historical data. It is worth noting in this context that from year to year, there is an improvement in the efficiency of microprocessor devices, an increase in data volumes, and an increase in the number of relevant information technologies (IT), which allows for optimizing the process of data analysis and monitoring changes in system parameters.

Among the methods of intelligent data analysis, we consider an approach using computer vision algorithms for more accurate segmentation of surface defects of various engineering objects. Automation of the process of diagnosing damage to the objects under investigation is carried out using a camera attached to an autonomous device. In the case of monitoring elements of constructions such as underground

pipelines, we consider surface defects (in particular, corrosion and crack type), for identification of which we use known implementations of neural networks. In order to optimize the process and training time of machine learning models, it was performed using an Nvidia GeForce RTX 3080 Ti graphics processor in the laboratory of artificial intelligence technologies of the Ivan Franko National University of Lviv.

The hardware helps to obtain information about the influence of the environment on changes in the current state of defects on the surface of structural elements (in particular, pipelines) using external sensors, taking into account communication between devices based on wireless data transmission technologies (Wi-Fi, Bluetooth, etc.). In this context, it is appropriate to use the methods and techniques of the Internet of Things (IoT) technology, as they allow transferring a set of collected data from sensors (for example: information about relative humidity and the number of counted pixels in the frame) to a central data store for further analysis.

Programming of diagnostic and monitoring systems for damage analysis (in particular, defects) requires the formation of several statistical models and methods for evaluating critical parameters that characterize the material condition of structural elements. Machine learning methods allow, based on flexible models, to automate the processes of classification and forecasting regarding changes in the state of defects. We check the forecast results based on the mentioned recognition methods using a Raspberry Pi microcomputer and the Nvidia JetRacer Kit.

From the analysis of recent research from open sources, it follows that the known amount of results for the analysis and prediction of the change in the state of defects of structural elements based on statistical models with the help of information technology (IT) and taking into account computer vision methods are not sufficient to be able to use them for processing of information about defects on the surface of engineering structures (in particular, underground pipelines). That is why it is worth applying methods of predictive analysis to process data read from sensors (cameras, humidity, temperature changes). The appropriate approach makes it possible to investigate the



effectiveness of a wide range of methods and techniques for the recognition of defect images and to analyze changes in the elements of engineering structures taking into account the current climatic conditions.

It should be noted that the dissertation work is devoted to the use of methods and means of intellectual analysis of changes in the state of surface defects in structural elements (in particular, in pipelines, and engineering objects). In this context, the dissertation also offers options for applying physical-mathematical, statistical, and probability modeling approaches for assessing the critical state of surface defects for the cyber-physical system "UMP - CP" (UMP - underground metal pipeline, CP - cathodic protection).

In the first chapter of the dissertation, the results of the use of information technologies for the recognition of defects in materials with the use of robotic tools are considered. The current state and aspects of defect recognition based on convolutional neural networks and edge filters are analyzed.

The actual problems of IT design are determined, taking into account the concepts of using computer vision methods to analyze images of defects in dark places (i.e. on the inner surface of pipes). A significant contribution to the development of approaches, methods, models and information technologies based on image processing in the context of flaw detection and analysis of the state of defects on the surface of structural elements is the introduction of modern tools and approaches for more accurate and efficient identification of defects and automation of the process of their analysis.

In the second chapter of the dissertation, a statistical analysis was carried out and a forecast of the change in the state of defects was made. To describe the main set of data on accidents and mechanical damage during the operation of pipelines, the reports of the administration on the safety of pipelines and the behavior of hazardous materials according to the information of the US Department of Transportation, as well as similar data obtained in Ukraine, were used. The main dependencies were determined and correlations were established regarding the parameters of influence on the failure of

structural elements, in particular, such as: external influences, age of the material, duration of operation and availability of repairs. A comparison of parametric and probabilistic methods for determining changes in the state of defects and evaluating the effectiveness and accuracy of the forecast is carried out. The use of the obtained results makes it possible to more comprehensively diagnose the change in the state of the structural element and optimize the anti-corrosion protection system of this object.

The third section of the dissertation is devoted to the hardware and software implementation of the IT prototype for the validation of the used methods of intelligent data analysis and evaluation of the change in the state of defects in engineering objects.

The development idea consists in the synthesis and research of defect recognition methods using the Nvidia Jetson Nano microcomputer and defect analysis based on statistical modeling of relevant parameters in the context of a predictive model. A review and analysis of wireless data transmission methods and their processing directly on a microcomputer to ensure the autonomy of the system was performed.

In the fourth chapter, an approach to the selection of analytical tools for physical-mathematical modeling of the quality characteristics of anti-corrosion protection of structures and the application of relevant results for IT development is proposed. The appropriate information technology is based on a physico-mathematical model of surface phenomena and methods of evaluating qualitative informational parameters. As a result of the implementation of IT, a new and improved comprehensive resource assessment option for "UMP-CP" system was developed in the context of computer modeling and takes into account optimization criteria, methods and techniques for evaluating changes in parameters and information flows.

**Keywords:** intellectual data analysis, machine learning, computer vision, recognition, Internet of Things, image mining, regression models, cyber-physical system, information flow, REST API, Raspberry PI, Nvidia Jetson Nano, cracks, pittings, defectoscopy.

## СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ

**Наукові праці, в яких опубліковані основні наукові результати дисертації**

### *Публікації у фахових виданнях України категорії Б:*

1. Шувар Р., Продивус А., Юзевич В., Огірко І., Огірко О., Ковтко Р., Мисюк Р. Інформаційні технології та загрози у кіберфізичних системах для відображення інформації у підземних металоконструкціях з дефектами // Штучний інтелект. – 2021. – Вип. 26 (1). – С. 85–94.
2. Mysiuk R., Yuzevych V., Mysiuk I. Data processing and visualization for corrosion detection // Electronics and Information Technologies. – 2021. – Vol. 16. – P. 50–60.
3. Mysiuk R., Yuzevych V., Mysiuk I., Ohirko I. Detecting cracks in concrete based on images using Amazon Web Service Rekognition // Electronics and Information Technologies. – 2022. – Vol. 20. – P. 18–29.
4. Mysiuk R., Yuzevych V., Mysiuk I. Api test automation of search functionality with artificial intelligence // Artificial Intelligence. – 2022. – Vol. 27, No 1. – P. 269–274.
5. Mysiuk R., Yuzevych V. IoT-based solution for detection defects in infrastructure objects using Raspberry PI // Electronics and Information Technologies. – 2023. – No. 21. – P. 45–56.

### *Публікації у виданнях, які проіндексовані у міжнародній наукометричній базі Scopus:*

1. R.V. Mysiuk, V.M. Yuzevych, M.F. Yasinskyi, S.V. Kniaz, Z.A. Duriagina, V.V. Kulyk. Determination of conditions for loss of bearing capacity of underground ammonia pipelines based on the monitoring data and flexible search algorithms // Archives of Materials Science and Engineering – 2022. – Vol. 115, No. 1. – P. 13–20.
2. High Availability System for Monitoring Material Degradation Processes at the Concrete-polymer Interface / R. Mysiuk, V. Yuzevych, B. Koman and M. Yasinskyi //

2022 12<sup>th</sup> International Conference on Advanced Computer Information Technologies (ACIT) – Spisska Kapitula, Slovakia, 26–28 September 2022. – P. 415–418.

3. Simulation of Corrosion Fracture of Nano-Concrete at the Interface with Reinforcement Taking into Account Temperature Change / R. Dzhalá, V. Yuzevych, R. Mysiuk, V. Brych, R. Skrynkovskyy, V. Lozovan, Y. Tyrkalo // CEUR Workshop Proceedings. MoMLLeT+DS 2022: 4<sup>th</sup> International Workshop on Modern Machine Learning Technologies and Data Science, November, 25–26, 2022, Leiden-Lviv, The Netherlands-Ukraine – Vol. 3312 – P. 123–133.

4. Video-based Concrete Road Damage Assessment Using JetRacer Kit / R. Mysiuk, I. Mysiuk, G. Pawlowski, V. Yuzevych, M. Yasynskiy and Y. Tyrkalo // 2023 17<sup>th</sup> International Conference on the Experience of Designing and Application of CAD Systems (CADSM), Jaroslaw, Poland, February 22–25, 2023. – P. 1–4.

5. Detection of Structure Changes in Lightweight Concrete with Nanoparticles Using Computer Vision Methods in the Construction Industry / R. Mysiuk, V. Yuzevych, B. Koman, Y. Tyrkalo, O. Farat, I. Mysiuk, L. Harasym // Proceedings of Eighth International Congress on Information and Communication Technology (ICICT), Lecture Notes in Networks and Systems, vol 694. Springer, Singapore. United Kingdom, London, 20–23 February 2023 – P. 339–348.

6. Detection of Surface Defects Inside Concrete Pipelines Using Trained Model on JetRacer Kit / Mysiuk R., Yuzevych V., Mysiuk I. Tyrkalo Y., Pavlenchuk A., Dalyk V. // 2023 IEEE 13th International Conference on Electronics and Information Technologies (ELIT), Lviv, Ukraine, September 26 – 28, 2023. – P. 21–24.

***Публікації у наукових періодичних виданнях інших держав:***

1. Mysiuk R., Mysiuk I., Pawlowski G. Data Modeling of Physical-Mechanical Processes in Nanoconcrete with the Ensemble of Pores // Path of Science. – 2022. – V. 8, No. 9 – P. 1001–1009.

2. Mysiuk R., Mysiuk I., Yuzevych V., Pawlowski G. Determining the Place of Depressurization of Underground Pipelines (Gas Pipelines): New Solutions in Industry based on Thermal Image Analysis Using Computer Vision // Path of Science. – 2022. – Vol. 8, No. 10 – P. 1001–1010.
3. Mysiuk R., Yuzevych V. Recover Data about Detected Defects of Underground Metal Elements of Constructions in Amazon Elasticsearch Service // Path of Science. – 2023. – Vol. 9, No. 1 – P. 1011–1019.
4. Mysiuk R. Pipeline Damage Assessment Based on Corrosion Segmentation Using JetRacer Kit // Path of Science. – 2023. – Vol. 9, No. 8 – P. 6001–6006.
5. Mysiuk R. Information technology for risk-based pipeline resource assessment // International scientific journal "Internauka". – 2024. – №1.
6. Obshta A., Yuzevych V., Pohrebniak A., Mysiuk R., Chorniy B. Diagnostics of oil leaks caused by malicious damage to the linear part of oil pipelines: innovative solutions for the oil industry // International scientific journal "Internauka". — 2024. — №2.
7. Mysiuk R. Towards Information Flows in Recognition and Prediction Tasks with Internet of Things // Path of Science. – 2024. – Vol. 10, No.1 – P.1001–1004.

***Наукові праці, які засвідчують апробацію матеріалів дисертації:***

1. Мисюк Р., Юзевич В. Система пошуку та відбору інформації про дефекти типу тріщин у базах знань // Міжнародна наукова інтернет-конференція “Інформаційне суспільство: технологічні, економічні та технічні аспекти становлення (випуск 53)” / Збірник тез доповідей: випуск 53 (м. Тернопіль, 16 листопада 2020 р.). – Частина 2. – Тернопіль. – 2020. – С. 57–60.
2. Mysiuk R., Yuzevych V., Koman B. Quality control of underground metal constructions with corrosion fatigue // Quality management in education and industry:

experience, problems and perspectives: proceedings of the V international scientific practical conference, May 20–21, 2021 / Lviv: «Piramida», 2021. P. 161–163

3. Мисюк Р., Юзевич В. Гнучкий алгоритм у системі пошуку та відбору інформації про дефекти типу тріщин // Міжнародна наукова інтернет-конференція “Інформаційне суспільство: технологічні, економічні та технічні аспекти становлення (випуск 55)” / Збірник тез доповідей: випуск 55 (м. Тернопіль, 9 лютого 2021 р.). – Тернопіль. – 2021. – С. 41–44.

4. Mysiuk R., Yuzevych V. Diagnosis of corrosion fatigue in underground metal constructions using genetic algorithms // III International Scientific and Practical Internet Conference “Mathematics and Informatics in Higher Education: Challenges of Modernity”, dedicated to the memory of Professors O. A. Pankov and V. S. Trokhymenko (Vinnytsia, May 20–21, 2021): book of abstracts. [Electronic network scientific publication], Vinnytsia, 2021, P. 93–96.

5. Mysiuk R. Statistical models of pipe configurations for assessment of defects in infrastructure objects // Information Technologies and Automation - 2023 / Proceedings of the XVI International Scientific and Practical Conference. Odesa, October 19–20, 2023. - Odesa, ONTU Publishing House – 2023 – P. 69–71.

6. Mysiuk R., Yuzevych V. Features importance in statistical models for detecting material cracks // Quality management in education and industry: experience, problems and perspectives: proceedings of the VI International Scientific and Practical Conference, November 16 – 17, 2023, - Lviv, Lviv Polytechnic Publishing House – 2023 – P. 152–153.

7. Mysiuk R., Mysiuk I., Yuzevych V., Shuvar R., Tsyuh S., Pavlenchuk N. Predictive Analysis of Macro Defects in Engineering Structures Using Machine Learning Technologies // International Conference on Optimization and Data Science in Industrial Engineering, November 16-17, 2023, Istinye University, Istanbul, Turkey – P.111.

## Зміст

ВСТУП.....	18
РОЗДІЛ 1. АНАЛІЗ МЕТОДІВ ТА ЗАСОБІВ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ЗОБРАЖЕНЬ ПОВЕРХНЕВИХ ДЕФЕКТІВ НА ПОВЕРХНЯХ У ЕІК.....	26
1.1. Методи та засоби інтелектуального аналізу даних.....	28
1.2. Концепція та актуальні проблеми ідентифікації дефектів.....	31
1.3. Методи машинного та глибокого навчання для діагностики поверхневих дефектів.....	35
1.4. Огляд систем діагностування ЕІК.....	39
1.5. Архітектура нейронної мережі для вирішення завдань діагностики на основі зображень.....	41
1.6. Особливості ідентифікації та обробки зображень дефектів на поверхнях ЕІК.....	49
1.6.1. Особливості ідентифікації та обробки зображень корозійних дефектів .....	49
1.6.2. Особливості обробки зображень ЕІК з пошкодженнями типу тріщин.....	53
1.6.3. Особливості обробки зображень ЕІК під час розгерметизації.....	55
1.7. Методи обробки зображень на мікрокомп'ютері в ІТ для оцінювання змін інформаційних потоків.....	58
Висновки до розділу 1.....	63
РОЗДІЛ 2. МЕТОДИ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ЗМІН СТАНІВ ДЕФЕКТІВ У ЕІК.....	66
2.1. Аналіз методів та етапів обробки даних про ЕІК.....	67
2.2. Опис та формалізація інформативних показників ЕІК.....	69
2.3. Статистичний аналіз інформативних параметрів.....	73
2.4. Аналіз залежностей між інформативними даними.....	77
2.5. Регресійний аналіз важливості ознак моделі ресурсу ЕІК.....	83
2.6. Формалізація станів дефектів на поверхнях.....	89
2.7. Методи оцінювання точності прогнозу змін станів поверхневих дефектів .....	90
Висновки до розділу 2.....	95

РОЗДІЛ 3. ЗАСОБИ ВАЛІДАЦІЇ ВИКОРИСТАНИХ МЕТОДІВ ДЛЯ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ЗМІНИ СТАНУ ДЕФЕКТІВ У ЕІК .....	97
3.1. Постановка задачі побудови системи розпізнавання для перевірки.....	97
3.2. Концепція прототипу інформаційно-діагностичної системи .....	98
3.3. Метод оптимізації передачі даних з IoT .....	101
3.4. Доповнення методу розпізнавання дефектів результатами прогностичного аналізу .....	105
3.5. Апаратна архітектура ІТ для оцінювання ресурсу кіберфізичної системи “МПТ – СКЗ” .....	106
3.6. Програмна архітектура для розпізнавання дефектів у кіберфізичній системі “МПТ – СКЗ” .....	109
3.6.1. Структура сервісної частини прогнозування для дослідження кіберфізичної системи “МПТ – СКЗ” .....	112
3.6.2. Програмні компоненти для дослідження кіберфізичної системи “МПТ – СКЗ” .....	118
3.7. Процес та результати валідації моделей .....	122
3.8. Переваги, обмеження та подальший розвиток ІТ .....	124
Висновки до розділу 3 .....	126
РОЗДІЛ 4. МЕТОДИ ОЦІНЮВАННЯ ДІАГНОСТИКИ ТА АНАЛІЗУ ІНФОРМАЦІЙНИХ ПОТОКІВ.....	128
4.1. Постановка задачі оцінки характеристик інформаційних потоків.....	128
4.2. Розв’язування та дослідження варіантів задачі оптимізації для оцінювання якості кіберфізичної системи “МПТ – СКЗ” .....	136
4.3. Статистична обробка даних для оцінювання критичного стану функціонування інженерного об’єкта з дефектами .....	143
4.4. Метод обчислення ефективних параметрів діагностики, які характеризують покращення якості кіберфізичної системи “МПТ – СКЗ” ...	148
Висновки до розділу 4 .....	151
ВИСНОВКИ.....	152
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	155
Додаток А .....	172



**ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ,  
СКОРОЧЕНЬ І ТЕРМІНІВ**

ЕК – елементи інженерних конструкцій

CNN – згорткова нейронна мережа

МН – машинне навчання

ГН – глибоке навчання

НМ – нейронна мережа

ІТ – інформаційна технологія

ІоТ – технології інтернету речей

МПТ – металевий підземний трубопровід

СКЗ – система катодного захисту

ЖЦ – життєвий цикл

## ВСТУП

**Актуальність теми.** З розвитком інформаційних технологій існує велика кількість інструментів для автоматизації процесу ідентифікації дефектів у інженерних об'єктах (зокрема, у елементах трубопроводів). Причиною виникнення відповідних пошкоджень матеріалів часто є зовнішні фактори, середовище та фактори постійної експлуатації. Застосування інформаційних технологій (ІТ) для діагностики систем з дефектами дозволяє забезпечити якість функціонування та оптимізацію роботи систем. Одним із способів оцінки пошкоджень на поверхнях є використання методів комп'ютерного зору та машинного навчання з використанням бібліотек (OpenCV, PyTorch, SimpleCV, Caffe, Tensorflow). Розпізнавання дефектів полягає в опрацюванні вхідного інформаційного відеопотоку та виділення місць виявлених пошкоджень на зображеннях. Враховуючи відповідні дані, можна здійснювати оцінювання критичних станів пошкоджень (тріщин, пітингів) матеріалів для локалізації знайдених місць їх дислокації на поверхнях. Автономні пристрої із вбудованою камерою дозволяють ефективніше здійснювати розпізнавання найбільш непомітних деталей на поверхнях елементів конструкцій.

Однією із важливих задач перед дослідниками є пошук найбільш ефективного методу діагностики об'єктів з дефектами у важкодоступних місцях на основі аналітики зображень. Для задач дефектоскопії реалізується постійне зчитування даних з сенсорів, а також регулярні перевірки інформації за допомогою приладів. Перспективним комерційним підходом є впровадження моніторингових систем на основі системи сенсорів. Серед українських розробок у галузі дефектоскопії на основі аналізу зображень варто звернути увагу на прилади Фізико-механічного інституту ім. Г.В. Карпенка НАН України для визначення 3D переміщень і деформацій поверхні. Одним з таких пристроїв є оптико-цифровий дефектоскоп. Крім цього, є портативні програмно-апаратні системи такі як: магнетоакустична система МАЕ - 1ЛР, ультразвуковий томограф

UST-05M та вібродіагностична система VECTOR. Вихідними даними для діагностування елементів конструкцій з дефектами є зібрані ультразвукові, магнетоакустичні та вібраційні дані з урахуванням підключення спеціальної техніки дротовим з'єднанням. Передача даних між компонентами сучасної інформаційної технології (IT) реалізується за допомогою бездротових технологій передачі даних (Wi-Fi, Bluetooth, LTE, Zigbee, LoRaWAN), які часто інтегровані з хмарними сервісами (Amazon Web Service, Microsoft Azure, Google Cloud Platform) для збереження та аналізу результатів.

Діагностика стану об'єктів на основі кіберфізичної системи враховує поєднання останніх розробок програмно-апаратних комплексів для збору, обробки та аналізу даних з фізичним світом. Проектування архітектури такого підходу базується на використанні сенсорів (температури, вологості та камери), портативних обчислювальних вузлів (мікроконтролерів чи мікрокомп'ютерів) та хмарних сервісів. Незважаючи на велику кількість реалізацій та пропозицій на ринку, методів розпізнавання пошкоджень об'єктів та проектування відповідних систем діагностики є недостатнім з погляду вибору найбільш оптимального рішення для оцінювання стану дефекту. Варто відзначити, що останніми роками розробка програмного забезпечення часто здійснюється з використанням елементів машинного навчання для покращення ефективності, точності та автоматизації процесів.

Функціонування інженерних об'єктів дозволяє забезпечувати транспортну логістику та основні потреби людини. Важливою проблемою є можливість реалізації процедури прогнозування зміни стану дефектів для забезпечення довговічності експлуатації елементів конструкцій. Використовуючи методи класифікації машинного навчання, можна здійснювати прогнозування найбільш ймовірних пошкоджень на поверхнях ЕІК.

Звернемо увагу на підбір параметрів, які найбільше впливають на зміни стану дефектів у ЕІК. Аналіз літературних джерел показав, що основні зовнішні

фактори впливу пов'язані з температурою, вологістю тощо, внутрішні (наявність антикорозійного захисту, тип речовини) та особливості форми об'єктів (довжина, товщина, тощо). Для цього у дисертаційній роботі проаналізовані причини аварій елементів конструкцій з урахуванням масиву характеристик матеріалу для формування наборів даних.

Для поточного стану досліджуваної проблематики та отриманих результатів залишається відкритою низка проблем. По-перше, це розробка методик моніторингу дефектів у елементах конструкцій методами комп'ютерного зору та виділення важливих особливостей їх поведінки. По-друге, застосування статистичних моделей для прогнозування пошкоджень у ЕІК. По-третє, застосування класифікаторів для прогнозу дефектів у ЕІК.

Аналіз сучасного стану робіт щодо дефектоскопії показав, що здебільшого для вирішення проблеми використання ІТ для інженерних об'єктів застосовують методи машинного навчання, які мають позитивні і негативні сторони. Слід звернути увагу на розробку комплексних підходів для перевірки результатів прогнозу на основі роботехнічних засобів.

**Зв'язок роботи із науковими програмами, планами, темами.** Дисертаційні дослідження виконувалися у Львівському національному університеті імені Івана Франка на кафедрі системного проектування в рамках науково-дослідних тем в межах робочого часу «Аналіз даних засобами машинного навчання» (номер державної реєстрації: 0119U002409, доц. Шувар Р. Я., 2021 р.) та «Системний аналіз даних засобами штучного інтелекту та машинного навчання» (номер держреєстрації: 0122U200688, доцент Шувар Р. Я., 2022–2024 рр.), які відповідають науковому напрямку розробки.

**Мета і задачі дослідження.** Метою дисертаційної роботи є розробка апаратно-програмного комплексу для оцінювання та прогнозування змін стану поверхневих дефектів на основі сформованих моделей у елементах інженерних конструкцій.

Для досягнення поставленої мети необхідно було розв'язати такі *задачі*:

1. Проаналізувати наявні інформаційні технології і статистичні моделі дефектоскопії з урахуванням поверхневих характеристик об'єктів;
2. Розробити функціональну структуру програмної реалізації методів для сегментації та локалізації тріщин та корозійних дефектів на поверхні елементів конструкцій на основі зображень з використанням методів комп'ютерного зору;
3. Розробити алгоритм роботи з даними, а саме процес відбору, збереження, опрацювання та аналізу даних про дефекти інженерних елементів конструкцій;
4. Побудувати та проаналізувати залежності впливу основних параметрів на основі зміни стану дефектів методами машинного навчання;
5. Проаналізувати взаємодію зібраних даних з сенсорів для перевірки прогнозу зміни стану дефектів у інженерних об'єктах та оцінити вплив інтенсивності інформаційних потоків у задачах розпізнавання та прогнозування;
6. Побудувати модель станів дефектів та визначити ефективні параметри для виконання діагностики кіберфізичної системи «МТП-СКЗ».

*Об'єктом дослідження* є процеси розпізнавання та прогнозування станів дефектів на поверхнях у елементах конструкцій (інженерних об'єктах).

*Предметом дослідження* є методи, моделі та алгоритми обробки даних (тобто елементи інформаційних технологій) для розпізнавання та прогнозування зміни станів дефектів на поверхнях у інженерних об'єктах.

*Методи дослідження.* У дисертаційній роботі для розв'язання поставлених завдань використані такі методи: для обробки та аналізу сенсорних даних інженерних об'єктів – методи математичної (регресійно-кореляційний аналіз) та описової статистики; для опрацювання інформації про кіберфізичну систему – морфологічна обробка зображень та підхід штучних нейронних мереж; для побудови фізико-математичних моделей – теорію дефектоскопії, статистичні методи; а під час розроблення програмних модулів – принципи об'єктно-

орієнтованого програмування, клієнт-серверної архітектури, REST, теорію нереляційних баз даних та технології EKL.

**Наукова новизна отриманих результатів.** На основі виконаних теоретично-практичних досліджень отримано наступні результати:

- вперше на основі регресійних моделей і низки критеріїв розроблено метод для перевірки якості прогнозу та оцінювання ресурсу елементів інженерних конструкцій за умови врахування зовнішніх впливів та інтенсивності інформаційних потоків в контексті інтелектуального аналізу даних про поверхневі дефекти;
- вперше в контексті інформаційних технологій розроблено і доведено до результативного варіанту модель відбору, опрацювання й аналізу даних про фактори впливу та параметри контактуючих середовищ, і в цьому контексті розв'язано оптимізаційні задачі обміну даних між компонентами системи, на основі яких прийнято рішення про ефективність інформаційних потоків, які є корисні для роботи з даними у сфері дефектоскопії;
- розроблено гібридну модель діагностичної системи для аналізу зміни стану тріщин та корозійних дефектів у різних елементах конструкцій за допомогою процедури ідентифікації зображень поверхневих пошкоджень елементів інженерних конструкцій (ЕІК) на основі аналізу потоку інформації про характеристики зображень;
- розроблено новий удосконалений комплексний варіант методу оцінювання ресурсу для кіберфізичної системи “металевий підземний трубопровід (МПП) – система катодного захисту (СКЗ)” у контексті оцінювання критичного стану дефектів на поверхнях розділу середовищ методами комп'ютерного зору;
- удосконалено функціональну структуру потоків даних у контексті інформаційної технології для опрацювання зображень дефектів (зокрема,

тепловізійних), а також для проведення оцінювання ресурсу елементів конструкцій на цій основі;

- отримав подальший розвиток спосіб представлення даних щодо ідентифікації поверхневих дефектів ЕІК на зображеннях у контексті обміну інформації, який базується на групуванні виділених пікселів, що характеризують ефективність інформаційних потоків;
- отримав подальший розвиток комплексний метод уточнення параметрів для ідентифікації критичних характеристик дефектів і на основі відповідних результатів проведено аналіз ресурсу елементів конструкцій.

**Практичне значення отриманих результатів** полягає в покращенні існуючих підходів для розпізнавання та прогнозування пошкоджень з урахуванням специфіки їх функціонування та факторів впливу на основі аналізу зображень поверхні інженерного об'єкта з використанням методів машинного навчання, а саме:

- розроблено регресійні моделі для оцінювання параметрів у сфері дефектоскопії, що дозволяють прогнозувати зміни стану дефектів інженерних об'єктів з урахуванням зовнішніх впливів;
- запропоновано фізико-статистичні моделі та методи опрацювання даних, що дають змогу безпосередньо визначати характеристики зміни стану дефектів об'єктів, що забезпечує більш гнучке налаштування діагностичних автоматизованих систем, зокрема, для елементів конструкцій, які перебувають у агресивному середовищі;
- реалізовано варіанти аналізу апаратно-програмної взаємодії в контексті інформаційних потоків з урахуванням IoT і методів передачі даних з обробленого зображення та сховища даних, що дозволяє вдосконалювати та покращувати процес розробки інформаційних систем для аналізу елементів конструкцій, а також приймати рішення щодо покращення режимів їх експлуатації;

- запропоновано методику опрацювання даних про дефекти з допомогою відеозображень і використано отримані результати для оцінювання критичного стану інженерного об'єкту на основі методів комп'ютерного зору;

Отримані у роботі результати впроваджені у відповідні навчальні курси у Львівському національному університеті імені Івана Франка.

**Особистий внесок здобувача.** Результати роботи, наведені в дисертації, отримано автором особисто. В рамках роботи над публікаціями здобувачеві належать такі результати: у працях [1, 2, 4, 5, 8, 18-21] – побудовані моделі для опрацювання й аналізу даних про поверхневі дефекти, у статтях [3, 7] – виконано аналіз зображень поверхневих дефектів із застосуванням методів комп'ютерного зору, у працях [6, 10, 11, 12, 14, 22-24] проаналізовані параметри впливу навантажень на параметри зміни стану дефектів для побудови відповідних моделей, у статтях [9, 13, 15] – алгоритм зі застосуванням методів комп'ютерного зору для аналізу даних на основі робототехнічних засобів, у статтях [16, 43] – доповнена модель оцінки ризиків та системи діагностики для дефектів на поверхнях на основі кіберфізичної системи металевий підземний трубопровід – система катодного захисту, у статті [17] – метод оптимізації інформаційні потоки даних для ефективною передачі інформації.

**Апробація результатів дисертації.** Основні положення та результати дисертації доповідались та обговорювались на: Міжнародній науковій інтернет-конференції “Інформаційне суспільство: технологічні, економічні та технічні аспекти становлення” (Тернопіль, 2020, 2021); XX та XXI Міжнародній науково-технічна конференція “Штучний інтелект та інтелектуальні системи” (International Scientific and Technical Conference “Artificial Intelligence and Intelligent systems”), (AIPS), (Київ, 2020, 2021); V та VI Міжнародна науково-практична конференція "Управління якістю в освіті та промисловості: досвід, проблеми та перспективи" (QM), (Львів, 2021, 2023); III Міжнародна науково-



практична конференція “Математика та інформатика в науці й освіті: виклики сучасності” (International Scientific and Practical Internet Conference “Mathematics and Informatics in Higher Education: Challenges of Modernity”), (Вінниця, 2021); 12<sup>th</sup> International Conference on Advanced Computer Information Technologies (ACIT) (Spisska Kapitula, Slovakia, 2022); 4<sup>th</sup> International Workshop on Modern Machine Learning Technologies and Data Science (MoMLeT+DS) (Leiden-Lviv, The Netherlands-Ukraine, 2022); 17<sup>th</sup> International Conference on the Experience of Designing and Application of CAD Systems (CADSM), (Jaroslaw, Poland, 2023); 8<sup>th</sup> International Congress on Information and Communication Technology (ICICT 2023), (London, United Kingdom, 2023); Міжнародна конференція 2023 IEEE 13th International Conference on Electronics and Information Technologies (ELIT), (Львів, 2023); XVI International Scientific and Practical Conference “Information Technologies and Automation - 2023” (ITIA), (Одеса, 2023); International Conference on Optimization and Data Science in Industrial Engineering (ODSIE 2023), (Istanbul, Turkey, 2023), а також на звітних конференціях Львівського національного університету імені Івана Франка за (2020-2024).

**Публікації.** За час написання дисертації опубліковано 25 наукових праць, серед яких 13 статей у наукових журналах та 12 публікацій у матеріалах конференцій. Серед публікацій 5 статей [1-5] опубліковано у журналах з переліку фахових видань відповідно до списку затверджених МОН України, 8 статей [6-10, 16, 17, 43] у міжнародних наукових журналах (стаття [10] включена до міжнародної наукометричної бази Scopus). Також видано 7 тез доповідей у матеріалах всеукраїнських та міжнародних наукових конференцій [18-24], а також 5 статей [11-15] у матеріалах конференцій, які реферуються у Scopus.

**Структура дисертації.** Дисертація складається із анотації, вступу, чотирьох розділів із висновками, списку використаних джерел (149 позиції) та додатку із публікаціями здобувача. Загальний обсяг становить 177 сторінок. Робота містить 60 рисунків та 12 таблиць.

## **РОЗДІЛ 1. АНАЛІЗ МЕТОДІВ ТА ЗАСОБІВ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ЗОБРАЖЕНЬ ПОВЕРХНЕВИХ ДЕФЕКТІВ НА ПОВЕРХНЯХ У ЕІК**

За даними американської технологічної дослідницької фірми Gartner [25] про нові технології та тренди відзначено, що протягом найближчих років у перспективі очікується значний розвиток методик граничного комп'ютерного зору (Edge Computer Vision). Поєднання підходів та методів дозволить більш точно розпізнавати об'єкти із застосуванням технологій штучного інтелекту і витратити при цьому менше матеріальних ресурсів. Аналіз об'єктів за допомогою роботехнічних засобів дозволяє автоматизувати відповідні процеси. Дослідники [26] показали можливості розпізнавання дефектів під час поєднання методик роботехнічних засобів та комп'ютерного зору. Відповідні підходи є досить популярними і дозволяють дослідити важкодоступні місця. Зростання популярних розробок та досліджень у напрямку розпізнавання дефектів впливає на інвестиційну привабливість систем з комп'ютерним зором та на можливість оцінювання технологічних ризиків, які є результатом впливу зовнішніх факторів на елементи інженерних конструкцій (ЕІК).

Особливістю обробки зображень під час діагностики ЕІК можуть бути розмиття та нечіткі частини зображень об'єктів. Пристрій зйомки може бути налаштований на різну фокусну відстань з передніми, основними та задніми об'єктами, які потрібно врахувати для уникнення розмиття. Наприклад, інформація про краї елементів, форму об'єкту, тінь чи текстуру допомагає розпізнавати на зображенні конкретний елемент досліджуваного дефекту.

Під час обробки зображень можливе потрапляння тіней з інших об'єктів чи залишків мастил у досліджувану область, і ці неточності (тіні) можуть впливати на коректність розпізнавання дефектів. Результатом аналізу зображення можуть бути фрагменти із структурними елементами, які потребують уточнення опису чи місць розташування дефектів. Відповідні практичні приклади, для яких варто

використовувати методи обробки низки зображень можуть бути пов'язані з корозійними дефектами у трубопроводах та процесами зміни тріщин на мостах. Особливо важливою відповідна інформація є під час розпізнавання та дослідження певних структурних елементів. Крім цього, дослідження у цьому напрямку можуть бути основою для побудови більш точних методів технічної діагностики об'єктів на основі аналізу зображень тріщин та пітингів.

Результатом обробки цифрових зображень є набір виділених фрагментів, які складаються з найдрібніших частинок (пікселів). Множини відповідних колірних елементів можуть характеризувати ступінь пошкодження і використовуватися для оцінки критичного стану дефектів. Важливим в контексті аналізу деталей з допомогою зображень є камера з високою роздільною здатністю, яка дозволяє здійснювати в інтеграції з мікрокомп'ютером обробку, збереження та надсилання результатів у сховище даних.

Одним із недостатньо досліджених аспектів інформаційних технологій (ІТ) є способи передачі даних опрацьованого зображення на основі технологій Інтернету речей. Запропоновано спрощений метод надсилання кількості пікселів для аналізу стану дефектів та визначення критичних місць. Зібрані дані можуть зберігатися локально або надсилатися у віддалене сховище даних для подальшого аналізу.

До актуальних проблем розробки ІТ у межах Інтернету речей належать: а) масштабованість та кількість даних, б) оптимізація кількості інформації для передачі та отримання, в) безпека та конфіденційність даних, г) складність системної розробки.

Описані проблеми є актуальними та потребують комплексного розв'язання для класу задач розпізнавання та аналізу дефектів у матеріалах з використанням ІТ.

## 1.1. Методи та засоби інтелектуального аналізу даних

Інтелектуальний аналіз даних (ІАД) є сукупністю методів та засобів, що застосовується для ідентифікації раніше невідомих залежностей та тенденцій змін значень параметрів чи отримання нової важливої інформації на основі великих масивів даних [27]. У процесі аналізу даних можна виділити наступні процеси роботи з даними: перетворення, очищення, візуалізацію та моделювання даних. Загалом, можна виділити кілька важливих частин інтелектуального аналізу даних (Табл 1.1): статистичне моделювання, машинне та глибоке навчання, генетичні алгоритми, обробка зображень та комп'ютерне зорове розпізнавання (CV), обробка часових рядів та гібридні підходи.

Таблиця 1.1 Переваги та недоліки методів інтелектуального аналізу даних для зміни стану дефектів на поверхнях

Методи та засоби ІАД		Переваги	Недоліки
Статистичне моделювання	<ul style="list-style-type: none"> <li>● Лінійна регресія</li> <li>● Логістична регресія</li> <li>● Аналіз варіації (ANOVA)</li> <li>● Кореляційний аналіз</li> </ul>	Дозволяє виявляти статистичні закономірності у даних та може використовуватися для визначення параметрів впливу на зміну станів дефектів на поверхні	Обмежена здатність моделювати складні зв'язки та нелінійність. Потребує чітко визначених припущень та модельних умов
Машинне та глибоке навчання	<ul style="list-style-type: none"> <li>● Метод опорних векторів (SVM) навчання</li> <li>● Випадковий ліс (Random Forest) - Нейронні</li> </ul>	Здатність робити складні залежності та класифікації	Вимагає великого обсягу даних для тренування моделей, можливе перенавчання

	<p>мережі (Neural Networks)</p> <ul style="list-style-type: none"> <li>• Згорткові нейронні мережі (CNN)</li> <li>• Рекурентні нейронні мережі (RNN)</li> <li>• Варіанти глибоких архітектур, такі як LSTM, GRU тощо</li> </ul>		
Генетичні алгоритми	<ul style="list-style-type: none"> <li>• Генетичний алгоритм (Genetic Algorithm)</li> <li>• Еволюційні стратегії (Evolutionary Strategies)</li> <li>• Генетичне програмування (Genetic Programming)</li> </ul>	Здатність пристосовуватися та знаходити оптимальні рішення у великих просторах параметрів	Вимагає налаштування параметрів, ретельного вибору функцій відліку
Обробка зображень та комп'ютерне зорове розпізнавання	<ul style="list-style-type: none"> <li>• OpenCV</li> <li>• TensorFlow</li> <li>• PyTorch</li> <li>• Keras</li> </ul>	Можливість автоматичного аналізу та розпізнавання об'єктів на зображеннях	Вимагає великого обсягу даних для тренування моделей, оптимального освітлення та зменшення шумів на зображеннях
Обробка часових рядів	<ul style="list-style-type: none"> <li>• Авторегресійні моделі</li> </ul>	Можливість прогнозування	Чутливість до шуму та аномалій

	(ARIMA) Експоненційн е згладжування (Exponential Smoothing) <ul style="list-style-type: none"> <li>• Методи              машинного              та              глибинного              навчання для              прогнозуванн              я часових              рядів</li> </ul>	змін у часі	у часових рядах
Гібридні підходи	<ul style="list-style-type: none"> <li>• Комбінація методів машинного та глибинного навчання зі статистичним аналізом</li> <li>• Ансамблеві методи (Ensemble Methods)</li> </ul>	Комбінація переваг різних методів для досягнення кращих результатів конкретних завданнях	Складність в налаштуванні та підтримці системи

Загалом, в контексті оцінювання точності та ефективності розглянутих методів ІАД мета аналізу даних може бути зведена до виборі функції  $f(\alpha)$  з найменшою похибкою

$$\min R(\alpha) = \frac{1}{n} \sum_{i=1}^n C[y_i, \alpha(x_i)] \quad (1.1)$$

де  $\alpha$  – можливі функцій, а  $C[y_i, \alpha(x_i)]$  – функція втрат із значенням залежної змінної  $\alpha(x_i)$  та відомі значення  $y_i$ .

Для бінарної класифікації, що визначає приналежність об'єкта до одного з двох класів найпростіша функція втрат є кількість правильно чи неправильно віднесених значення до класу 0 або 1. Для багатокласової класифікації, коли

об'єкт може належати до будь-якого з кількох класів, функція втрат стає складнішою через необхідність урахування різних типів помилок класифікації. В такому випадку для кожного можливого співпадіння прогнозованого та істинного класів створюється матриця втрат розміром  $k \times k$ , де  $k$  - це кількість класів.

Для регресійних методів зазвичай використовується середньоквадратична помилка (Mean Squared Error, MSE) або середня абсолютна помилка (Mean Absolute Error, MAE) як функція втрат.

## **1.2. Концепція та актуальні проблеми ідентифікації дефектів**

Забезпечення точності під час діагностування дефектів є важливим аспектом безпеки експлуатації об'єктів і елементів конструкцій. При цьому існують різні способи дефектоскопії: візуальний, акустичної, магнітної, електростатична тощо. Точність кожного з методів залежить від низки параметрів та чинників впливу. Найбільший вплив на руйнування об'єкта спостерігається на етапах збільшення розмірів дефектів до більш помітних критичних значень на поверхнях елементів конструкцій. Відповідну групу дефектів називають поверхневими.

З результатів досліджень у роботі [28] можна зробити висновок, що набувають популярності методи комп'ютерного або машинного зору для ідентифікації дефектів у режимі реального часу. Кожен з методів обробки зображення з використанням нейронних мереж має низку переваг та недоліків. Наприклад, у роботах [29, 30] один варіант потребує більше уваги для попередньої обробки зображення, а другий – розставлення міток на картині зображення. Натомість, процедура фільтрування не потребує навчання, а з допомогою нейронних мереж можна більш точно визначати місця пошкоджень.

Останніми роками особливо інтенсивно використовують технології штучного інтелекту для аналізу дефектів у матеріалах, а саме глибоке навчання

та машинне навчання [31]. Однією із проблем для покращення розпізнавання об'єктів є наявність підготовлених великих наборів даних необхідних під час навчання нейронних мереж. При цьому виділяють наступні групи підходів до навчання: навчання з учителем (контрольовані), напівконтрольовані та навчання без учителя (неконтрольовані). З одного боку, рівень ідентифікації дефектів у навчанні з учителем може бути дуже високим, враховуючи наявність позначених зразків зображень з дефектами та без них.

Натомість може виникати дисбаланс класів у наборах даних. З іншого боку, при навчанні без учителя можна автоматично класифікувати вхідні дані через вивчення зв'язків та характеристик без прив'язок до міток. Поєднанням обох попередніх методів навчання є напівконтрольований. Перевагою можна вважати використання меншого набору даних так, як у процесі навчання встановлюються умови відсутності дефектів. Точність такого методу є зазвичай меншою за навчання з вчителем.

В табл. 1.2 наведено назви та посилання на набори даних для елементів конструкцій.

Таблиця 1.2. Інформація про набори даних з поверхневими дефектами для деяких матеріалів

Назва набору даних та література	Тип матеріалу	Посилання
Severstal: Steel Defect Detection	Сталь	<a href="https://kaggle.com/c/severstal-steel-defect-detection/data">https://kaggle.com/c/severstal-steel-defect-detection/data</a> (перевірено 13 листопада 2023 року)
GC10-DET Metallic Surface Defect Detection	Метал	<a href="https://kaggle.com/alex000kim/gc10-det">https://kaggle.com/alex000kim/gc10-det</a>



		(перевірено 13 листопада 2023 року)
Industrial Metallic Surface Dataset	Метал	<a href="https://kaggle.com/datasets/ujik132016/industrial-metallic-surface-dataset">https://kaggle.com/datasets/ujik132016/industrial-metallic-surface-dataset</a> (перевірено 13 листопада 2023 року)
Structural Defects Network (SDNET)	Бетон	<a href="https://kaggle.com/datasets/aniruddhsharma/structural-defects-network-concrete-crack-images">https://kaggle.com/datasets/aniruddhsharma/structural-defects-network-concrete-crack-images</a> (перевірено 13 листопада 2023 року)
Concrete Defect Bridge Image Dataset	Бетон	<a href="https://zenodo.org/record/2620293#.Y1dPO3bMKUk">https://zenodo.org/record/2620293#.Y1dPO3bMKUk</a> (перевірено 13 листопада 2023 року)
Surface Crack Detection Dataset	Бетон	<a href="https://kaggle.com/arunrk7/surface-crack-detection">https://kaggle.com/arunrk7/surface-crack-detection</a> (перевірено 13 листопада 2023 року)
CrackForest-dataset	Бетон	<a href="https://github.com/cuilimeng/CrackForest-dataset/tree/master/image">https://github.com/cuilimeng/CrackForest-dataset/tree/master/image</a>

concreteIn_inpection_VGGF		<a href="https://github.com/CCNYRoboticsLab/concreteIn_inpection_VGGF">https://github.com/CCNYRoboticsLab/concreteIn_inpection_VGGF</a>
DeepCrack	Бетон	<a href="https://github.com/yhlleo/DeepCrack/blob/master/dataset/DeepCrack.zip">https://github.com/yhlleo/DeepCrack/blob/master/dataset/DeepCrack.zip</a>
Bridge Cracks	Бетон з металом	<a href="https://github.com/Iskysir/Bridge_Crack_Image_Data">https://github.com/Iskysir/Bridge_Crack_Image_Data</a> (перевірено 13 листопада 2023 року)
Corrosion Computer Vision Project	Металом	<a href="https://universe.roboflow.com/khaingwintz-gmail-com/dataset--2-pathein-train-plus-v-3-update-mm">https://universe.roboflow.com/khaingwintz-gmail-com/dataset--2-pathein-train-plus-v-3-update-mm</a> (перевірено 13 листопада 2023 року)
MVTec AD	Різні матеріали	<a href="http://mvttec.com/company/research/datasets">http://mvttec.com/company/research/datasets</a> (перевірено 13 листопада 2023 року)

Набори даних певної тематики підбирались відповідно до поставленого завдання. Для забезпечення кращої точності моделі застосовано методи аугментації зображень у наборах даних за допомогою випадкових поворотів та зсувів.

### **1.3. Методи машинного та глибокого навчання для діагностики поверхневих дефектів**

Протягом останніх років застосування методів машинного та глибокого навчання для ідентифікації зображень значно зросло. Відповідні підходи використовуються для дефектоскопії. Відомі різні методи нейронних мереж з модифікаціями, які використовуються для розпізнавання поверхневих дефектів: CNN (Convolutional Neural Network) спеціалізується на обробці введення, яке має просторову структуру, таку як зображення, GAN (Generative Adversarial Network) складається з генератора та дискримінатора, мультиспостереження стерео (Stereoscopic Multi-Observation SDD) використовує дві або більше камери для отримання стерео зображення для отримання глибинної інформації, ResNet (Residual Neural Network) використовує "блоки зі зливанням" для полегшення навчання ГМ, Faster R-CNN (Faster Region-Based Convolutional Neural Network) включає RPN (Region Proposal Network) для вибору областей зображення, які ймовірно містять корисну інформацію, YOLO (You Only Look Once) визначає елементи зображень одним проходом нейронної мережі замість роздільної процедури пропозицій та класифікації.

Вибір конкретної архітектури часто залежить від конкретних вимог завдання, обсягу даних, швидкодії та інших факторів. Наприклад, у роботі [28, 32] проведений порівняльний аналіз результатів досліджень, у якому на основі аналізу точності класифікації та з урахуванням функції втрат мережі типу CNN встановлено варіанти ефективності для ідентифікації дефектів. Проведено аналогічні дослідження для архітектури інших типів. У багатьох випадках комбінація декількох методів або моделей може покращити результат.

Працездатність наведених методів переважно демонструється на різних матеріалах, або елементах конструкцій з досліджуваних джерел [29–35], як у Таблиці 1.3.

Таблиця 1.3. Останні аспекти досліджень з використанням машинного навчання та нейронних мереж для поверхневих дефектів у матеріалах

Автори	Використані методи	Класифікація при навчанні	Матеріал	Результати
Natarajan, V.; Hung, T.Y.; Vaikundam, S.; Chia, L.T. [34]	CNN and SVM	з учителем	Метал	У класифікації цей метод перевершує інші, оскільки враховано найсучасніші традиційні функції ручної роботи, а також певні глибокі функції ConvNet, отримані з попередньо вибраного найкращого шару на основі розгляду кількох наборів даних, в яких відображено інформацію про аномалії і текстури.
Fu, G.; Sun, P.; Zhu, W.; Yang, J.; Cao, Y.;	CNN	з учителем	Сталь	Відповідний підхід може забезпечити кращу точність класифікації ніж традиційні.

Yang, M.Y.; Cao, Y. [35]				
Di, H.; Ke, X.; Peng, Z.; Dongdong, Z. [36]	GAN	без учителя	Сталь	CAESGAN забезпечує найкращий рівень класифікації порівняно з традиційними методами, особливо для гарячокатаного листового прокату
Akhyar, F.; Lin, C.Y.; Mughtar, K.; Wu, T.Y.; Ng, H.F. [37]	SDD і ResNet	з учителем	Сталь	Ідентифікацію дефектів сталеві поверхні можна виконувати з високою швидкістю та заданою точністю
Li, K.; Wang, X.; Ji, L. [38]	Faster- RCNN	з учителем	Сталь	Досягнуто кращої точності ідентифікації зображень та більш точного визначення характеристик дефектів, особливо для дрібних і тонких

				дефектів
Cha, Y.- J.; Choi, W.; Buyukozt urk, O. [39]	CNN	з учителем	Бетон	Натренована нейронна мережа на 40000 зображень (256 × 256 пікселів). Досягнена точність приблизно 98%
Zhen Yu [40]	Yolo v5	напівконтрол ьований	Бетон	3500 зображень з них 500 з мітками тестові дані. Точність приблизно 84%
Khumaidi , A.; Yuniarno, E.M.; Purnomo, M.H. [41]	CNN	з учителем	Зварні шви	Точність розпізнавання 95.83%

Аналіз дефектів на основі зображень є комплексним процесом, що містить декілька споріднених задач:

- збір даних, який полягає у зчитуванні з приєднаної камери до діагностичного пристрою потоку вхідних зображень;
- опрацювання зображень, використовуючи низку перетворень з використанням методів машинного навчання;
- аналіз результатів обробки зображень з відображенням і виділенням областей пошкоджень.

#### 1.4. Огляд систем діагностування ЕІК

Виявлення та локалізацію дефектів хвильовим методом описано в роботі [42]. У цьому документі досить детально описано процес виявлення дефектів шляхом відбиття направлених хвиль від досліджуваних поверхонь. Для кожного типу дефекту існують різні методи та відповідно технології для ідентифікації змін станів. Першим етапом роботи з даними у діагностичних системах є процес збору та передачі даних (зокрема, з Modbus) для подальшого аналізу [43, 44].

Ще одним з способів проведення діагностики для виявлення витоків в нафтових трубопроводах відбувається у кіберфізичній системі на основі пристрою FPGA у праці [45]. При цьому основним параметром використано метод обчислення хвиль негативного тиску. Проте вищезгадані методи діагностики більше застосовуються для оцінки зміни конкретних параметрів і не можуть забезпечувати повноцінний аналіз дефектів. У багатьох випадках комбінація обох методів або використання їх у поєднанні з іншими методами може бути найбільш ефективним рішенням.

Комп'ютерний зір є потужним інструментом для ідентифікації дефектів на поверхнях, який забезпечує високу чутливість, точність та ефективність у широкому спектрі застосувань. Засобами дефектоскопії дослідники отримують різні рішення стосовно аналізу одного і того ж знімка. Останнім часом проводяться інтенсивні дослідження, спрямовані на розробку ефективних методів автоматичного розпізнавання дефектів.

Камера є основним сенсором для валідації прогнозованих дефектів на поверхні об'єкта, проте часто для схожих завдань використовують ендоскоп. Вибір між камерою та ендоскопом повинен здійснюватися враховуючи конкретні завдання, вимоги та умови застосування. Камера може бути оптимальним вибором для загального використання, тоді як ендоскоп може бути важливим для спеціалізованих процедур у вузьких місцях трубопроводів. Цей метод має переваги під час обробки зображень у порожнині важкодоступного місця в

об'єктах (трубопроводи) за рахунок мікровідеокамери, яка не здатна передавати дрібні деталі і мати спотворення у процесі сканування.

На ринку існує багато роботехнічних систем та платформ для розпізнавання дефектів у матеріалах. Дослідники аналізують у працях [27, 46, 47] та використовують наявні роботехнічні системи для ідентифікації дефектів у важкодоступних місцях інженерних об'єктів. В цьому контексті варто звернути увагу на Redzone Robotics з багатосенсорною інспекцією підземних ділянок. Така інспекція стосується роботи з ширококутною камерою, для якої характерний кут  $190^\circ$ . Відповідні дослідження характерні для США і Катару, з урахуванням прототипу роботи з сенсором Lidar та камерою.

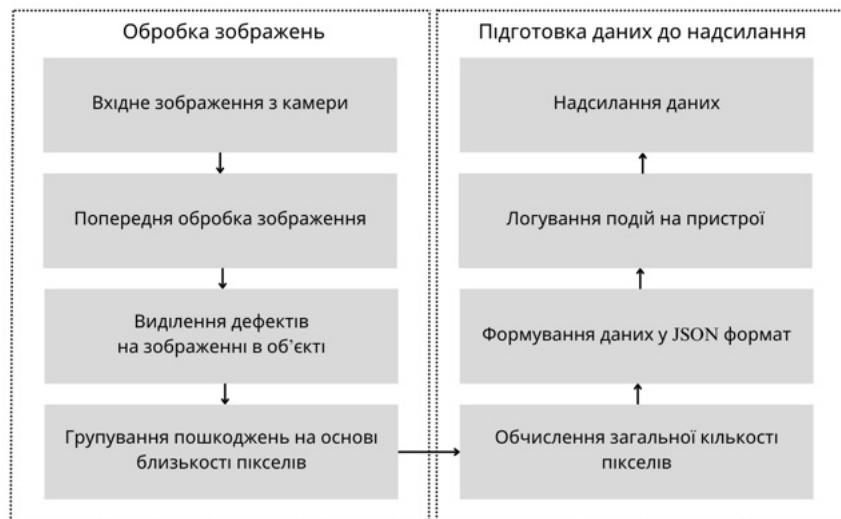


Рис. 1.1. Загальна структура процесу автоматизованого розпізнавання та підготовка даних для надсилання на реєстраційні пристрої

Процес автоматичного розпізнавання дефектів можна представити у виді послідовності чотирьох основних процедур (рис. 1.1): попередня обробка зображення, виділення дефектів на зображеннях об'єкта, групування пошкоджень на основі близькості пікселів, обчислення загальної кількості пікселів, формування даних у JSON формат, логування подій на пристрої, надсилання інформації до сховища даних.



Для ідентифікації дефектів у матеріалах важливу роль відіграє освітлення. Особливо це стосується виробничих процесів та якості контролю. Вплив освітлення може бути визначений різними способами з урахуванням відповідних факторів, таких як: тип світла, яскравість, кут падіння світла та колірна температура. Приклад розробки системи освітлення усередині труби подано у праці [27], у якому є інформація про використання кількох штучних джерел освітлення для рівномірного розподілення його по досліджуваній ділянці.

### **1.5. Архітектура нейронної мережі для вирішення завдань діагностики на основі зображень**

Для розпізнавання дефектів у матеріалах використовують різноманітні алгоритми та моделі обробки даних. Ці методи можна поділити на дві групи: класичні методи обробки зображень та сучасні методи глибокого навчання, які використовують нейронні мережі. Складність розробки методів обробки зображень у галузі дефектоскопії полягає у великій кількості різних дефектів та матеріалів, умов освітлення, наявності великої кількості навчальних даних та виборі оптимальних алгоритмів з оцінкою ефективності та надійності. Для розв'язку задач аналізу дефектів часто використовують синтез обох груп методів обробки зображень. Серед класичних методів можна виділити такі варіанти: фільтри зображень використовуються для виокремлення характерних особливостей з урахуванням розмиття для видалення шуму; морфологічні операції використовуються для обробки форми зображень; алгоритми ідентифікації контурів та порогування зображень використовуються для розділення областей на зображеннях, які мають значення пікселів вище або нижче заданого порогу. Результати досліджень, отримані у працях [31–41], показали інтенсивне використання сучасних методів глибокого навчання:

- Convolutional Neural Networks (CNN) для ідентифікації важливих особливостей на зображеннях та виявлення взаємозв'язків між ними.

- YOLO (You Only Look Once) модель для ідентифікації зображень у режимі реального часу.
- Region-based CNN (R-CNN) спеціалізовані для ідентифікації об'єктів та їх границь на зображеннях.
- Mask R-CNN є розширенням R-CNN з ідентифікацією зображень та виділенням відповідних областей на зображенні.
- Generative Adversarial Networks (GAN) використовують для генерації штучних прикладів дефектів для покращення тренувального набору.

Враховуючи вищезазначені особливості застосування таких алгоритмів для порівняння ефективності та точності у задачах розпізнавання дефектів на поверхнях на зображення реалізовано моделі CNN та YOLO. CNN - це загальна архітектура нейронної мережі для обробки зображень, в той час як YOLO - це конкретний алгоритм ідентифікації об'єктів, який використовує CNN для цієї задачі.

Під час реалізації CNN у задачах розпізнавання для сегментації зображень застосовують архітектуру глибокого навчання UNet. Така архітектура особливо ефективна для завдань, де необхідно визначити границі та форму елементів на зображенні розпізнавання дефектів [30]. Архітектура UNet складається з двох основних частин: енкодера (зменшує розмір зображення та визначає обсяг інформації) та декодера (збільшує розмір та відновлює вихідну інформацію). Основна ідея UNet полягає в тому, щоб використовувати інформацію з енкодера для точної сегментації об'єктів на зображенні. При цьому, кожен рівень енкодера з'єднується з відповідним рівнем декодера, створюючи “шлях еквівалентності”.

Щодо використання комп'ютерного зору для розпізнавання дефектів у пошкоджених елементах конструкцій, то така процедура може реалізовуватися через обробку зображень для ідентифікації аномалій або пошуку невідповідностей у текстурі, кольорі, а також формі матеріалу.

UNet дозволяє більш точно сегментувати зображення дефектів на основі використаних серій згорткових та з'єднуючих шарів (див. рис. 1.2). Таким чином, описана модель може класифікувати кожен піксель виконуючи поступове зменшення розміру зображення та збільшення глибини. На зворотньому шляху UNet (декодер) використано серію транспонованих згорткових шарів у поєднанні із звичайними згортковими шарами, щоб збільшити масштаб зображення для ідентифікації дефекту на поверхні ЕІК.

Певна кількість енкодерів та декодерів може змінюватися відповідно до різних конфігурацій та параметрів для конкретного набору даних. Наступний крок – оцінювання ефективності інформації на зображеннях дефектів за допомогою метрик якості і вибір тих даних, що дають найкращі результати. Тому оптимальні значення параметрів можна отримати експериментальним шляхом, зокрема, оцінити: глибину мережі (кількість шарів), роздільну здатність (розмір зображення), кількість фільтрів (каналів) на кожному шарі, функції активації (наприклад, ReLU) та функції втрат (наприклад, бінарна перехрестна ентропія для задачі бінарної сегментації), результати регуляризації (наприклад, dropout) та аугментації даних, налаштування гіперпараметрів.

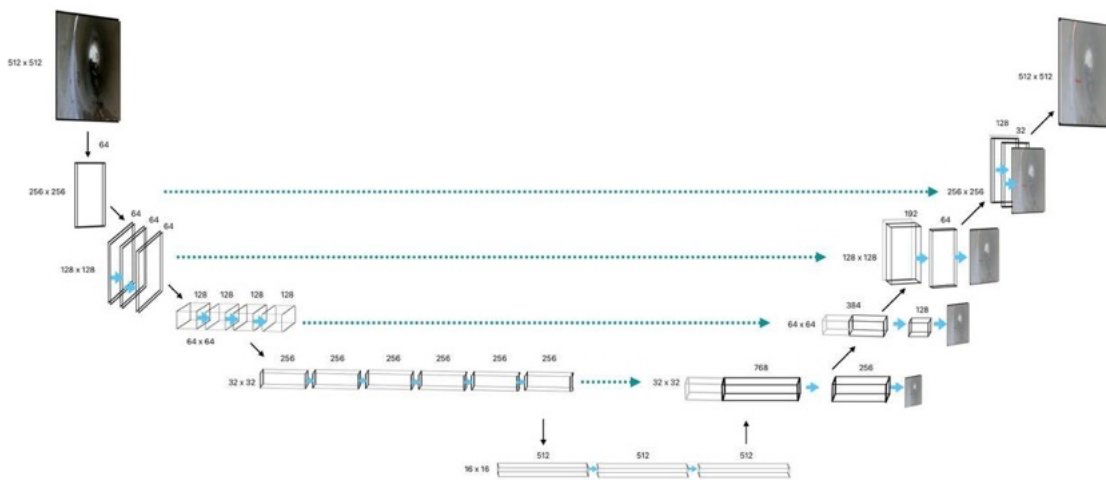


Рис. 1.2. Приклад UNet діаграми для CNN моделі

Створена модель CNN на основі UNet архітектури, яка навчена за допомогою бібліотеки PyTorch [48] та перетворено у формат ONNX. Зазвичай ONNX Runtime може значно сприяти у виконанні моделей машинного навчання, збережених у форматі ONNX, на різних пристроях та платформах, забезпечуючи високу продуктивність та ефективне використання ресурсів [49]. Тому для ще однієї реалізації використано бібліотеки ONNX Runtime та OpenCV для виконання оцінювання пошкоджень поверхонь ЕІК у режимі реальному часі. Архітектура відповідної моделі показана на Рис. 1.2.

Серед існуючих програмних реалізацій [50] сегментації дефектів з CNN використано набір даних для навчання (11200 стандартизованих зображень з розміром 448x448). Відповідна програма використана для порівняння з створеними моделями CNN та YOLO, а результати ефективності наведені у Табл. 1.3.

Для наборів даних входять зображення з наступними типами дефектів на поверхнях (див. таблицю 1.1): присутні лише тріщини, схожі на дефекти типу тріщин (бруківка та плитка зі швами), тріщина із рослинами, тріщини з шумом на горизонтальних чи вертикальних поверхнях (стіна та бетон), інші предмети.

Для порівняння існуючий набір даних доповнено зображеннями дефектів типу тріщин на циліндричних об'єктах (включно з проведеною аугментацією пов'язаною з поворотами). У роботі використано архітектуру ГН VGG16 (Visual Geometry Group 16), що складається з 16 шарів (включаючи 13 конволюційних і 3 повнозв'язаних шарів). Хоча ResNet зазвичай складається з більшої кількості шарів (наприклад, ResNet-50 має 50 шарів) і гарантує більшу точність, проте VGG16 є більш простішим в контексті архітектури [51].

Навчання цієї реалізації нейронної мережі відбувалося у лабораторії технологій штучного інтелекту Львівського національного університету імені Івана Франка на графічному процесорі Nvidia GeForce RTX 3080 Ti з об'ємом пам'яті 12 Гб. При цьому на обчислюваних ресурсах встановлена операційна

система Ubuntu 22.04.2 LTS, мова програмування Python 3.10.12 та основними бібліотеками Pytorch версії 2.2.1, torchvision версії 0.17.1 та onnxoptimizer версії 0.3.13.

В режимі реального часу для ідентифікації об'єктів виконано ще одну реалізацію глибокого навчання з допомогою бібліотеки YOLO [52]. Основна ідея цією програмної бібліотеки полягає в тому, щоб поділити зображення на сітку та визначити області, де знаходяться об'єкти, а також їх клас із зазначеною впевненістю. Першим етапом такої реалізації є підготовка даних для навчання, тому усі зображення аугментовані з застосування випадкових поворотів (+/- 5%) та поділені на навчальні (70%), валідаційні (20%) та тестові (10%) набори даних. Розмір зображень змінено аналогічно як для CNN. У кожному зображенні для обчислення точності здійснено виділення очікуваних полігональних областей дефектів на поверхнях як на Рис.1.3. Roboflow [53] дозволяє здійснювати анотування та інтеграцію з іншими популярними інструментами машинного навчання, такими як TensorFlow, PyTorch, YOLO, Darknet та іншими.

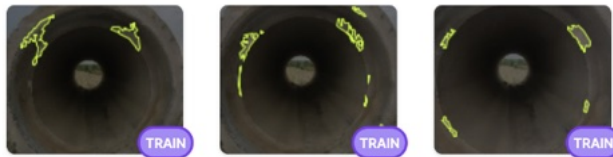


Рис. 1.3. Приклад підготовки даних для навчання у Roboflow

Навчання моделі відбулося за допомогою бібліотеки Ultralytics YOLOv8.0.196 [54] на мові програмування Python версії 3.10.12 та використовує бібліотеку PyTorch версії 2.1.0 з підтримкою CUDA (Compute Unified Device Architecture) версії 12.2. Для відповідної реалізації використано середовище Google Colab Research для прискорення обчислень на відповідному графічному прискорювачі NVIDIA Tesla T4 з обсягом пам'яті 15102MiB.

	from	n	params	module	arguments
0					
1	-1	1	928	ultralytics.nn.modules.conv.Conv	[3, 32, 3, 2]
2	-1	1	18560	ultralytics.nn.modules.conv.Conv	[32, 64, 3, 2]
3	-1	1	29056	ultralytics.nn.modules.block.C2f	[64, 64, 1, True]
4	-1	1	73984	ultralytics.nn.modules.conv.Conv	[64, 128, 3, 2]
5	-1	2	197632	ultralytics.nn.modules.block.C2f	[128, 128, 2, True]
6	-1	1	295424	ultralytics.nn.modules.conv.Conv	[128, 256, 3, 2]
7	-1	2	788480	ultralytics.nn.modules.block.C2f	[256, 256, 2, True]
8	-1	1	1180672	ultralytics.nn.modules.conv.Conv	[256, 512, 3, 2]
9	-1	1	1838080	ultralytics.nn.modules.block.C2f	[512, 512, 1, True]
10	-1	1	656896	ultralytics.nn.modules.block.SPPF	[512, 512, 5]
11	-1	1	0	torch.nn.modules.upsampling.Upsample	[None, 2, 'nearest']
12	[-1, 6]	1	0	ultralytics.nn.modules.conv.Concat	[1]
13	-1	1	591360	ultralytics.nn.modules.block.C2f	[768, 256, 1]
14	-1	1	0	torch.nn.modules.upsampling.Upsample	[None, 2, 'nearest']
15	[-1, 4]	1	0	ultralytics.nn.modules.conv.Concat	[1]
16	-1	1	148224	ultralytics.nn.modules.block.C2f	[384, 128, 1]
17	-1	1	147712	ultralytics.nn.modules.conv.Conv	[128, 128, 3, 2]
18	[-1, 12]	1	0	ultralytics.nn.modules.conv.Concat	[1]
19	-1	1	493056	ultralytics.nn.modules.block.C2f	[384, 256, 1]
20	-1	1	590336	ultralytics.nn.modules.conv.Conv	[256, 256, 3, 2]
21	[-1, 9]	1	0	ultralytics.nn.modules.conv.Concat	[1]
22	-1	1	1969152	ultralytics.nn.modules.block.C2f	[768, 512, 1]
23	[15, 18, 21]	1	2116435	ultralytics.nn.modules.head.Detect	[1, [128, 256, 512]]

Model summary: 225 layers, 11135987 parameters, 11135971 gradients, 28.6 GFLOPs

Рис. 1.4. Параметри архітектури нейронної мережі YOLO на основі попередньо навченої моделі

Навчання тривало приблизно 1 хвилину для 25 епох. Ця модель має 168 шари і загально 11,125,971 параметрів для навчання. Загальний час обробки одного зображення складається з часу попередньої обробки, визначення розміщення дефектів та додавання візуальних міток на зображення, які в середньому складають 3.5 мс, 18.2 мс і 28.3 мс відповідно.

Приклади результатів обробки зображень показані на рис.1.5 - рис. 1.9.

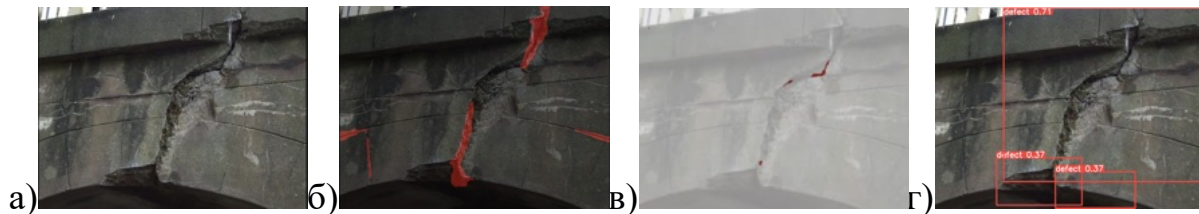


Рис.1.5. Приклад зображення для ідентифікації на основі мостової конструкції:

а) - вхідне, б) - оброблене [50], в) створена CNN, г) YOLO

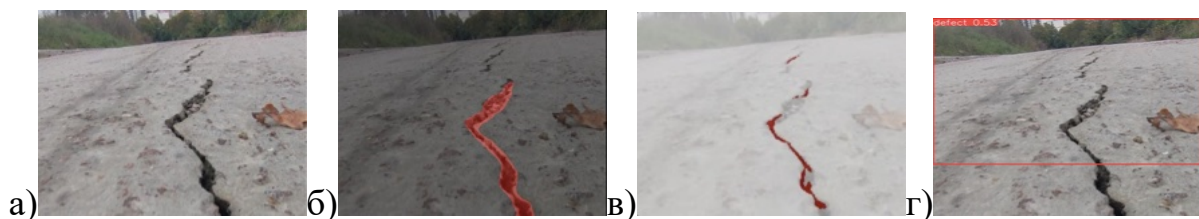


Рис.1.6. Приклад зображення для ідентифікації на основі горизонтальної конструкції: а) - вхідне, б) - оброблене [50], в) створена CNN, г) YOLO

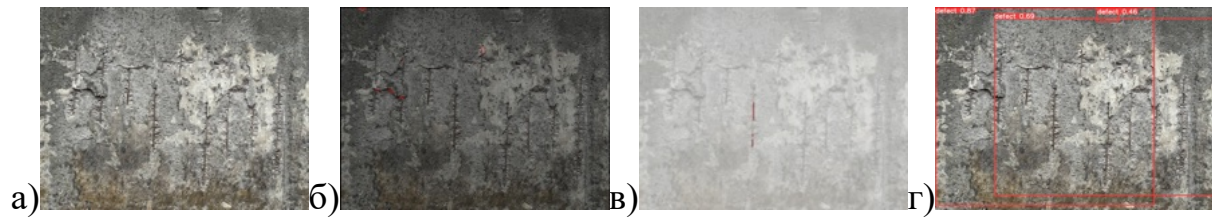


Рис.1.7. Приклад зображення для ідентифікації на основі вертикальної конструкції: а) - вхідне, б) - оброблене [50], в) створена CNN, г) YOLO

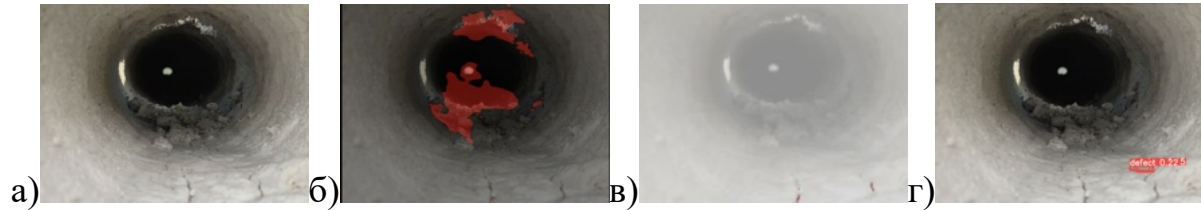


Рис.1.8. Приклад зображення для ідентифікації на основі конструкції циліндричної форми: а) - вхідне, б) - оброблене [50], в) створена CNN, г) YOLO

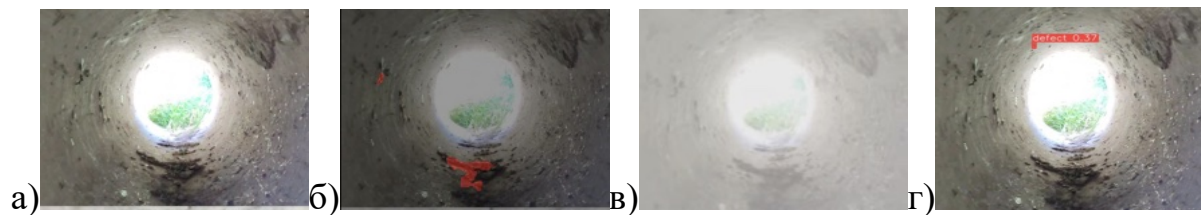


Рис.1.9. Приклад зображення для ідентифікації на основі конструкції циліндричної форми: а) - вхідне, б) - оброблене [50], в) створена CNN, г) YOLO

Особливістю візуалізації результату на зображеннях з YOLO є обмежувальні рамки, які позначають ідентифіковані об'єкти, використовуючи вектор значень

$$C_{1,1} = (P_c, B_x, B_y, B_w, B_h, C_1, C_2) \quad (1.2)$$

де,  $P_c$  – ймовірність прогнозу,  $B_x, B_y$  – центр рамки,  $B_w, B_h$  - ширина та висота рамки,  $C_1, C_2$  – відповідність класу (дефект чи не дефект).

Intersection over Union (IoU) та Dice є метриками для оцінки точності сегментації (розмежування) областей на зображеннях. У Табл. 1.4 ці параметри ефективності використовуються для оцінки точності моделей.

IoU площі об'єднання (область, що містить як вихідну, так і прогнозовану область) до площі перетину цих областей і обчислюється

$$IoU = \frac{TP}{TP+FP+FN} \quad (1.3)$$

де TP (True Positive) - кількість пікселів, які модель правильно визначила як область об'єкту; FP (False Positive) - кількість пікселів, які модель помилково визначила як область об'єкту; FN (False Negative) - кількість пікселів, які модель пропустила (не визначила як область об'єкту).

У контексті сегментації зображень, Dice coefficient також використовують як міру подібності між прогнозованою та справжньою масками і обчислюється за формулою

$$Dice = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \quad (1.4)$$

Таблиця 1.4. Показники якості сегментації під час ідентифікації дефектів

Реалізації ГН	середнє IoU	середнє Dice
CNN [50]	0.4687	0.6033
Навчена CNN	0.6327	0.7245
YOLO	0.5628	0.6823

У працях [55, 56] використано об'єднання кількох методів обробки зображень, що показує переваги у точності для ідентифікації об'єктів гібридних підходів у проектуванні систем.

Наведені вище методи можуть бути використані окремо або в комбінації, залежно від конкретних вимог задачі та наявності даних для тренування. Зазвичай вибір конкретного методу обробки даних залежить від обсягу та якості доступних даних, тому розширення наборів даних дозволило покращити розпізнавання дефектів типу тріщин у циліндричних формах (трубопроводах).



## **1.6. Особливості ідентифікації та обробки зображень дефектів на поверхнях ЕІК**

Серед діагностичних задач розглядають ідентифікацію, класифікацію та розпізнавання фрагментів на зображеннях, як це описано у праці [57]. Відтак, процес починається з аналізу, де для об'єктів або матеріалів визначають наявність аномалій. Такого типу задачі розглянуто у [58] з урахуванням наявності або відсутності корозії у металах за допомогою глибокого навчання з конкретним відсотком впевненості. Наступним етапом є класифікація, де ідентифіковані аномалії групуються за спільними характеристиками. Під час аналізу характеристик зображень дефектів варто враховувати: розмір, форму та приналежність їх до конкретного класу чи категорії. Наприклад, у праці [59] проведено класифікацію приналежності зображень дефектів до певної групи за ознаками певних стадій корозії: низька (менше 5% пікселів з виявленою корозією), середня (менше 15 % пікселів з виявленою корозією), висока (більше 15 % пікселів з виявленою корозією). Найбільш складним є етап розпізнавання, в якому визначають конкретний тип чи природу виявлених аномалій. Процес аналізу зображень полягає у використанні алгоритмів машинного навчання або інтелектуальних систем, згідно з яким автоматично визначають природу дефекту та його можливі зміни.

Оскільки в загальному зображення з дефектами елементів конструкцій мають доволі різноманітну структуру (пітинги [60], тріщини тощо), розглянемо в наступному підрозділі актуальні підходи до розпізнавання фрагментів зображень дефектів.

### **1.6.1. Особливості ідентифікації та обробки зображень корозійних дефектів**

Візуальною характеристикою корозії є колір та форма утворень (дефектів) на поверхні металів. Основні кольорові відтінки, характерні для корозії, можуть відрізнитися і залежать від типу матеріалу, а також від ступеня окислення

приповехневої фази. Ступінь критичного стану (стійкості) елемента конструкції можна визначити, застосувавши метод комп'ютерного зору для аналізу поверхневих дефектів [59]. Дефекти можна класифікувати, виділивши умовно три групи їх зображень, які відповідають критичній, середній та легкій корозії у навчальних наборах даних [47]. Проте такий метод потребує виконання навчання і може бути пов'язаний з додатковими налаштуваннями гіперпараметрів.

У праці [61] обґрунтовано вплив попередньої обробки на продуктивність деяких традиційних моделей обробки зображень фонових об'єктів з урахуванням колірних ефектів. Відповідними обмеженнями методу можна вважати: складність вибору оптимального порогового значення для бінаризації, необхідність видалення фону та загальна модифікація яскравості.

Визначення інформації про пошкоджені пікселі базується на принципі обчислення відсоткового співвідношення, в чисельнику якого знаходиться умовна площа покритих корозією ділянок, а в знаменнику – роздільна здатність обробленого зображення, аналогічно як у праці [62].

$$\text{пошкодження} = \frac{\text{Кількість пікселів ідентифікованого дефекту}}{\text{Загальна кількість пікселів зображення}} \quad (1.5)$$

З погляду ефективного проектування апаратних засобів для оцінювання якості проєктованих пристроїв проведено огляд систем і їх складових від найпростіших на основі Arduino Uno до різноманітних роботів “павуків”, які здатні рухатися та розпізнавати корозію (тобто ідентифікувати зображення дефектів) на внутрішній поверхні труб [63, 64].

Оскільки трубопроводи виготовлені зі сталі, то розглянемо можливості визначення найбільш пошкоджених місць на основі піксельного аналізу зображення (рис. 1.10).

На противагу використанню оцінки стану дефекту на поверхні ЕІК як незмінному (статичному) об'єкта використано набір зображень, що є послідовністю кадрів відео і показано зміни станів ідентифікованих пошкоджень у режимі реального часу.

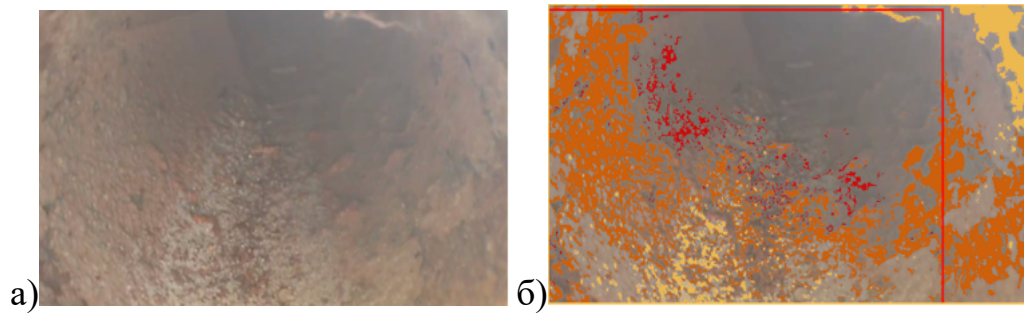


Рис. 1.10. Відеокадр з допомогою яких виявлено корозійні пошкодження в трубопроводі: а – вхідний, б – оброблений

Для опису та відтворення кольорів кожного дискретного елемента растрового зображення у комп'ютерній графіці використовують різні представлення RGB (Red, Green, Blue), CMYK (Cyan, Magenta, Yellow, Key/Black), HSV (Hue, Saturation, Value) та інші.

Порогові значення параметрів, які відповідають кольоровим характеристикам, можуть бути обрані для відокремлення певного фрагмента зображення, покритого корозією. У праці [9] досліджено і встановлено, що з використанням умов насиченості кольору можна відфільтрувати і виділити найбільш пошкоджені (окиснені) місця на зображеннях та ідентифікувати наявність критичних місць (дефектів) (рис. 1.11а). Для відокремлення інших рівнів пошкодження (середнього, легкого), це можна зробити, як показано на рис. 1.11б і рис. 1.11в відповідно. Ступені корозії у відсотках за рівнями пошкодженості виявилася за формулою (1.5) такими: високою - 1,535%, середньою - 21,357%, низькою - 4,01%.

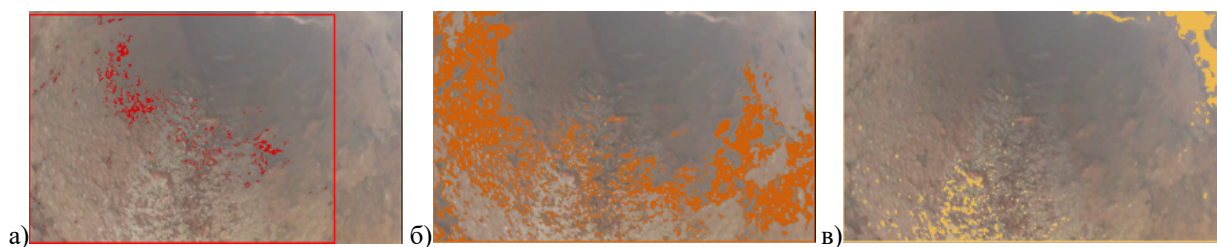


Рис. 1.11. Оброблений відеокадр з розділеною корозією за станом: а – критичний, б – середній, в – низький

Аналіз статистичних даних стосовно зображень дефектів дає можливість виявити місця на ділянці трубопроводу, які потребують термінового ремонту (наприклад, інформація на рис. 1.12а). Для побудови залежностей на рис. 1.12а використано інформацію про гістограми частот кольорових пікселів. Уточнення місця максимального корозійного ураження (зокрема, на рис. 1.12а цій ситуації відповідає фрейм з номером 15 (*frame number*)) пов'язано з відстеженням відеокадрів об'єктів [9].

Використовуючи гістограми з відповідними значеннями RGB, також можна визначити найвищу частоту кольорів, як показано на рис. 1.12б. Помітні вершини на гістограмі характеризують кількість пікселів, які відповідають певному кольору. Такого типу методики, з допомогою яких побудовано рис. 1.5а і рис. 1.12б, можна використовувати для уточнення умов зміни корозійного стану елементів конструкцій і враховувати при цьому результати аналізу відеокадрів [9].

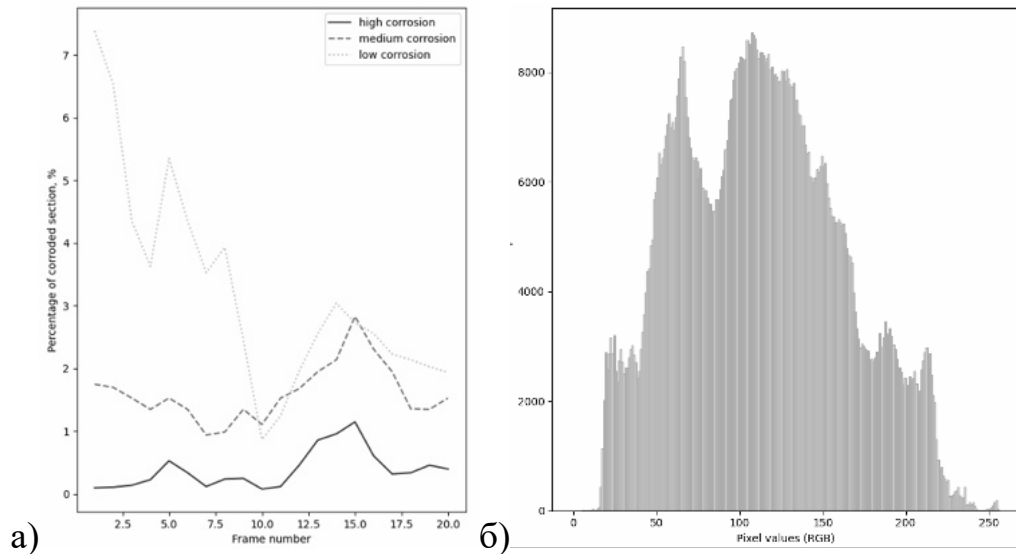


Рис. 1.12. Результати аналізу потоку зображень: а) – зміна безрозмірного параметра, який характеризує у відсотках ступінь корозії, в залежності від номера ділянки для металевого об'єкта; б) – гістограма частоти кольорових пікселів

### 1.6.2. Особливості обробки зображень ЕІК з пошкодженнями типу тріщин

Тріщини, як один з можливих дефектів у елементах конструкцій, можуть бути ідентифіковані у матеріалах на вертикальних (наприклад, на стінах чи підпірних об'єктах), горизонтальних (наприклад, на дорогах), циліндричних (наприклад, трубопроводи) площинах. Відповідно до цього, варто розглядати різні підходи до розпізнавання дефектів з використанням фільтрів або інших більш складних нейронних мереж.

Серед найбільш ефективних методів фільтрування зображень для виділення контурів об'єктів є Canny, Собеля та Робертса оператори, як описано у праці [65]. Результати відповідного дослідження демонструють, що Canny має найкращі показники з-поміж контурних фільтрів по безперервності, забезпечення позиціонування краю та ефекту сегментації. У праці [66] більшу перевагу надають використанню Canny у порівнянні з вищезгаданими зважаючи на більшу ефективність та точність. Його використання у комбінації з іншими фільтрами може бути ще більш ефективним для поверхневого оцінювання зміни стану доріг чи бетонних конструкцій [5]. Тому при використанні фільтрування зображення це дає перевагу у відсутності додаткових дій, а саме процесу навчання нейронних мереж [13].

Порівняно з горизонтальними чи вертикальними площинами у циліндричних за формою елементах конструкцій (рис.1.13а) локалізація дефекту може бути обмеженою просторовими можливостями камер чи інших засобів вимірювань. Тому в таких умовах ідентифікація дефекту є складнішим завданням.

Проте здебільшого у працях [27, 33–41] дослідники надають перевагу використанню нейронних мереж у питаннях ідентифікації дефектів у елементах конструкцій, ототожнюючи цей підхід з найбільш ефективним натепер.

З рис. 1.13в можна переконатися в тому, що більш точним для складного завдання з виявленням різних типів об'єктів вважається CNN. Проте варто звернути увагу що, при цьому цей метод потребує достатньої кількості даних для

навчання та більше часу на його опрацювання, особливо для глибоких мереж. Фільтрування за допомогою виділення контурів (наприклад, Canny) можна застосувати для миттєвого отримання результату та виділення дефектів на поверхні (рис.1.13г). Хоча при умові низької контрастності або наявності шуму на зображенні таке використання фільтрів може бути не достатньо ефективним порівняно з нейронними мережами, оскільки вимагає налаштування параметрів для досягнення оптимальних результатів на різних типах зображень.

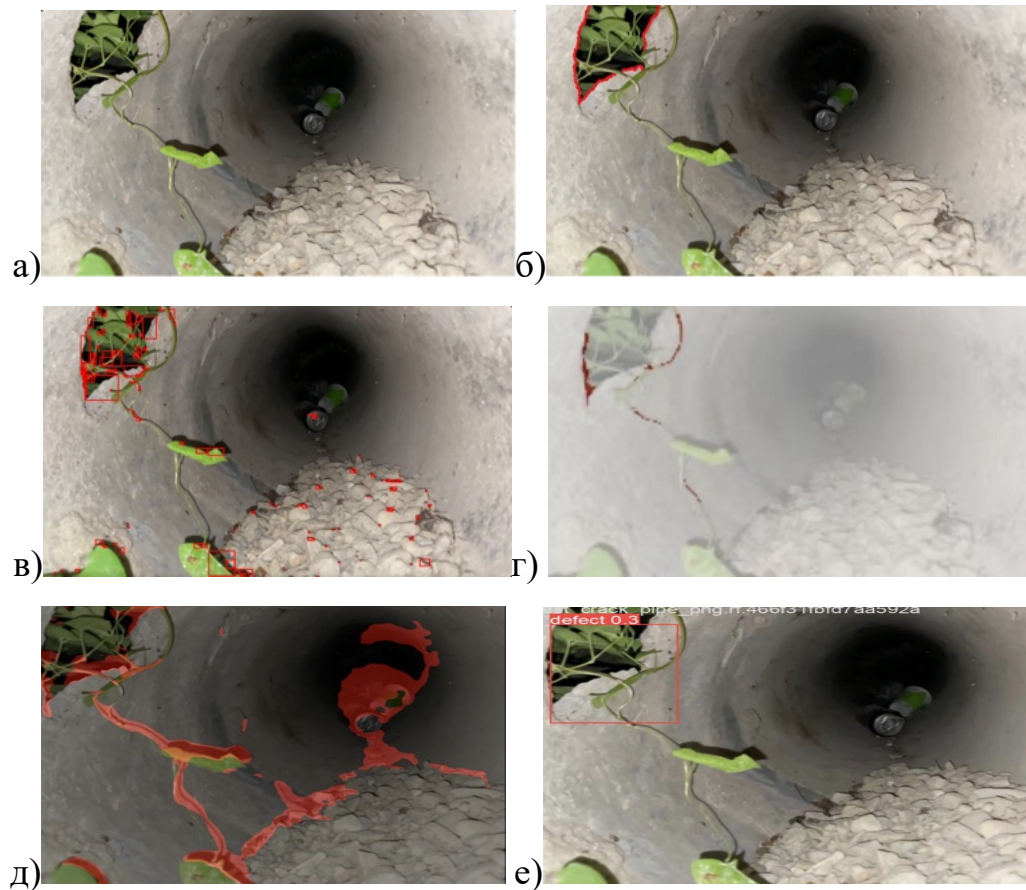


Рис. 1.13. Приклади обробки зображення з використанням: а) – вхідного, б) – очікуване, в) – фільтру Canny, г) – CNN, д) до нормалізації з YOLO, е) після нормалізації з YOLO

Одним із складних завдань є виявлення дефектів у закритих та неосвітлених частинах об'єкта таких як трубопроводи, яке потребує додаткового штучного освітлення [15]. Всередині яких можуть виникати дефекти, зміну стану

поверхні яких можна відслідковувати за допомогою оцінки поверхні методами комп'ютерного зору. Проте штучне освітлення може створювати перешкоди для розпізнавання такі як, тіні чи затемнення дефектів, що впливає на ефективність розпізнавання. Питання рівномірності освітлення розглянуте у праці [27], у якій пропонується застосувати кілька джерел освітлення. Відповідний підхід враховує особливості відбивання світла від поверхонь досліджуваного об'єкта.

### **1.6.3. Особливості обробки зображень ЕІК під час розгерметизації**

Процес розгерметизації у ЕІК вказує на втрату герметичності, тобто виникнення витoku рідини або газу через тріщини чи інші пошкодження стінок труби. Причинами розгерметизації можуть бути механічні ушкодження, корозія, зношення або неправильна експлуатація.

Тепловий контроль використовується для виявлення розгерметизації шляхом виявлення змін температури на певній ділянці пошкодженого ЕІК. Теплові камери або інші теплові датчики можуть виявити зони з аномальною температурою, що може свідчити про місце розгерметизації та уникнення техногенних катастроф.

Тепловий контроль може здійснюватися методами інфрачервоної дефектоскопії [67]. Тепловізійні зображення містять опис колірних характеристик об'єкта таким чином більш точно дозволяють оцінити потенційно небезпечні ділянки з дефектами. Основним сенсором для отримання таких типів зображень є тепловізійна камера чи плати розширення (зокрема, MLX90640-D110, Thermal Camera HAT, FLIR Lepton 3.5 IR Camera [68] чи інші аналоги). Інтеграція відповідних компонент до пристроїв обробки даних відбувається через з'єднувачі, які забезпечують обмін даних.

Спектр кольорів отриманого зображення відповідно до температури дає змогу оцінити від найтепліших (зазвичай позначені червоним або білим кольорами) до найхолодніших місць (зазвичай позначені синім або чорним

кольорами) та проаналізувати можливу розгерметизацію об'єктів відповідно. Оскільки колірні характеристики є структурними елементами кожного пікселя, то можуть використовуватися для оцінювання зображення. Аналогічно як і для обробки зображень для інших типів дефектів можна застосувати колірну сегментацію як запропоновано у праці [7]. Виділення певних колірних діапазонів дозволяє визначати потенційно небезпечні ділянки об'єкту та групувати пікселі за відстанню чи певними критеріями (наприклад, критичний стан дефекту чи типом дефекту).

На тепловізійному зображенні промислового елементу трубопроводу (рис. 1.14а), окрім місця пошкодження, ідентифіковано кілька ділянок, які можуть спричинити розгерметизацію металеві труби. Відсоток пошкодження ЕІК (N) на основі зображення може бути обчислений за формулою (1.1). Наприклад, зображення (рис. 1.14б) містить 0.13 % вмісту з ідентифікованими дефектами, що може використовуватися для аналізу критичного стану ЕІК.

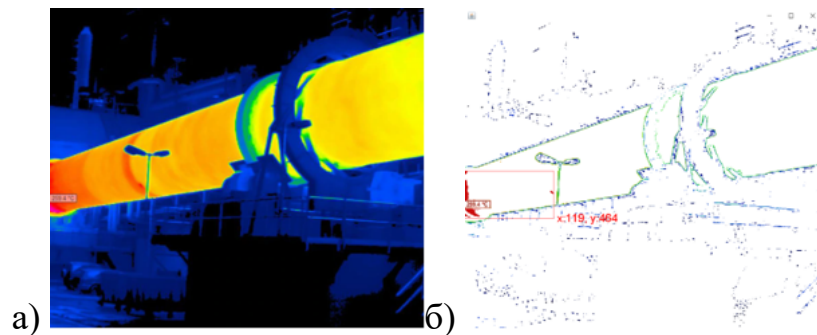


Рис. 1.14. Трубопровід на промисловому об'єкті: а – тепловізійне зображення трубопроводу, б – оброблене зображення трубопроводу.

Ще один приклад ЕІК під час розгерметизації показаний на рис. 1.15а, а пошкодження ідентифіковане на рис. 1.15б. У цьому випадку відсоток пошкодженої (червоної) області до загального розміру зображення становить 0,81.



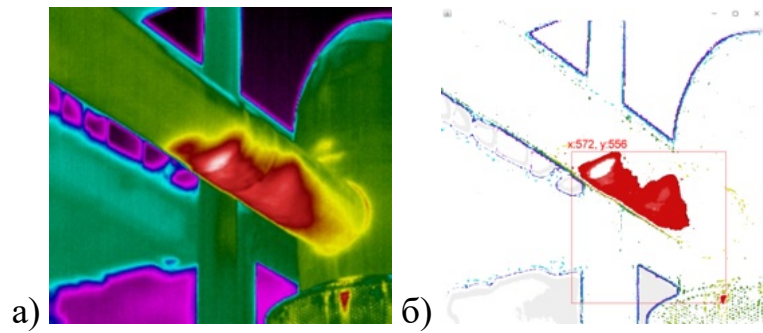


Рис. 1.15. Аналіз пошкодженого трубопроводу: а – термозображення трубопроводу, б – оброблене зображення трубопроводу.

Ділянка виділена білим кольором на тепловізійному позначає найвищу температуру і відповідно основну вершину дефекту під час розгерметизації. При цьому виділена група кольорів відтінків червоного кольору навколо вершини показує площу пошкодженої ділянки. З візуалізації відповідного типу дефекта на тепловізійному зображенні можна визначити непомітні деталі методом виділення кольору за насиченістю (Рис. 1.16).

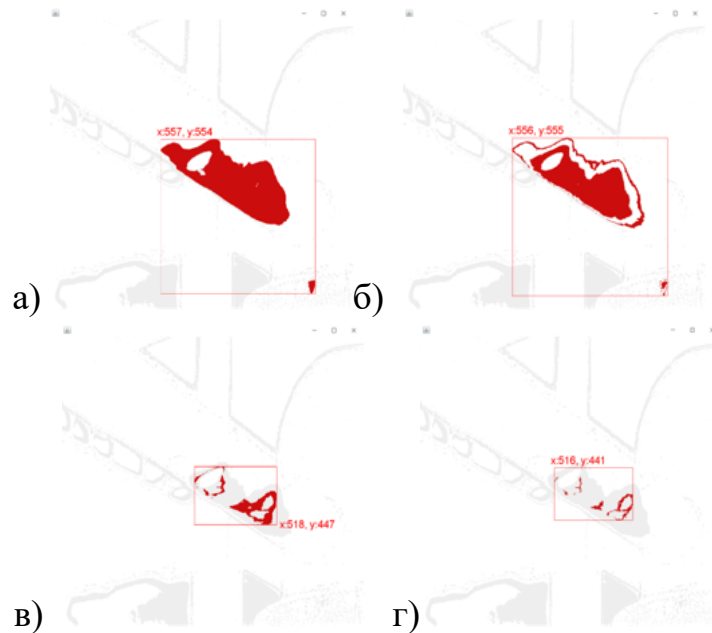


Рис. 1.16. Аналіз пошкодженого трубопроводу з різними відтінками кольору пошкодження: а – червоний = 100 і зелений = 70 і синій = 70, б – червоний = 150 і зелений = 100 і синій = 100, в - червоний = 200 і зелений = 50 і синій = 50, г - червоний = 210 і зелений = 50 і синій = 50

На рисунку 1.16а виділена площа пошкодження під час розгерметизації є достатньо великою і дорівнює 5,6% від розміру зображення.

Під час виділення кольору на RGB(150, 100, 100) спостерігається зменшення виділеної зони пошкодження до 4,3% від розміру зображення, а при RGB (200, 50, 50) - 0,8% від розміру зображення. Контури потенційного пошкодження візуально краще виділяються на малюнку 1.16г для комбінації RGB(210, 50, 50). У цьому випадку площа пошкодження становить 0,246% розміру зображення.

Таким чином, запропонований метод уточнення пошкодженої області ЕІК під час розгерметизації на основі насиченості кольору тепловізійного зображення для виділення контуру при вершині дефекту.

### **1.7. Методи обробки зображень на мікрокомп'ютері в ІТ для оцінювання змін інформаційних потоків**

Процес аналізу дефектів відбувається на основі вхідних даних, які потребують використання пристроїв відбору інформації. Крім цього, ще однією особливістю таких засобів є розміри та автономність обчислювального пристрою. Такими пристроями є мікрокомп'ютери (рис. 1.11). Характеристики мікрокомп'ютера впливають на ефективність та результативність ідентифікації дефектів на поверхнях ЕІК в контексті технологій Інтернету речей (ІоТ). Деякі моделі мікрокомп'ютерів мають обмежені ресурси, такі як обсяг оперативної пам'яті, потужність процесора та простір для зберігання даних.

Серед найбільш популярних мікрокомп'ютерів є Raspberry Pi та Nvidia Jetson Nano. Основною відмінністю між ними є специфіка використання. Наприклад, перший з перелічених використовується для реалізації різноманітних прикладних проектів, тоді як інший спеціалізується на завданнях пов'язаних з штучним інтелектом. При цьому такі пристрої можуть бути інтегровані з камерою та здійснювати аналіз потоку зображень в режимі реального часу.

Одним із показників оцінювання ефективності інформаційних потоків є обчислення часу обробки одного кадру в секунду (Frame per second) [69].

$$FPS = \frac{C}{T_{S2} - T_{S1}}, \quad (1.6)$$

де  $T_{S1}$  - початковий час,  $T_{S2}$  - кінцевий час,  $C$  - кількість кадрів за час відео

Для цього реалізований клас *FramePerSeconds* з методами *elapsed()* повертає загальну кількість секунд між початковим і кінцевим інтервалом та *framePerSeconds()* обчислюється (приблизну) кількість кадрів за секунду (FPS).

Кількість кадрів для відео в одну секунду 32, кількість кадрів з відео на 9 секунд - 273, на 20 секунд - 606. Приклад такого розкадрування відео показано на рис. 1.17. При обробці зображень з великою роздільною здатністю на малопотужних пристроях можуть бути певні затримки.

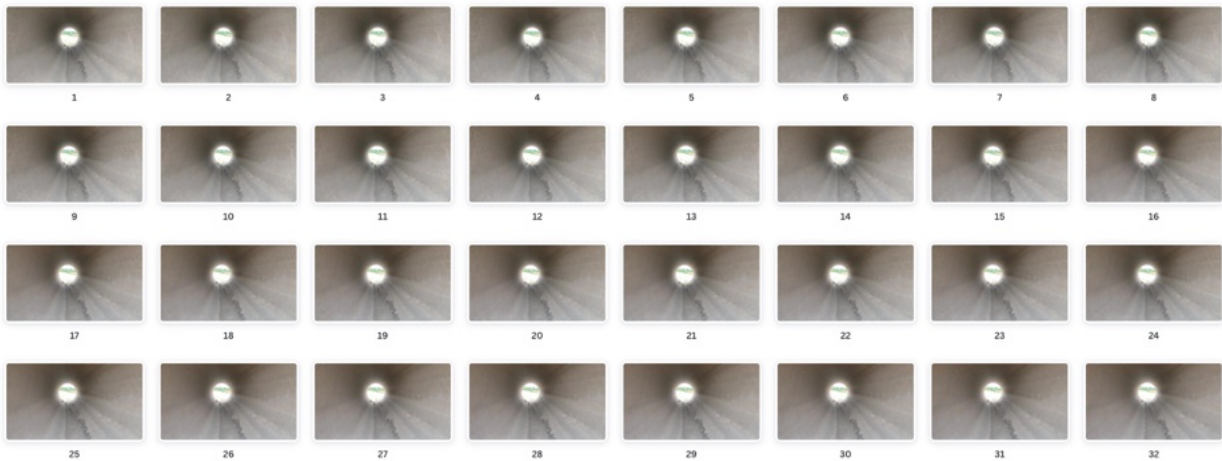


Рис. 1.17. Поділ відео 1 секунди на кадри

Оптимізацію моделі нейронних мереж можна здійснити з механізмом TensorRT, який підтримується у Nvidia Jetson Nano і використовується для роботи з навченими моделями на основі Caffe, TensorFlow, PyTorch і ONNX. Цей механізм дозволяє здійснювати оптимізацію пам'яті та пропускну здатності графічного процесора на основі даних про шари та процесів тензорного злиття. Таким чином, це зменшує накладні витрати на читання та запис тензорних даних для кожного шару. Така оптимізація на основі TensorRT дозволяє ідентифікувати дефекти на поверхнях у десятки разів швидше та з меншим часу відгуку.

Створена модель за допомогою PyTorch перетворюється у формат ONNX та після цього конвертується в файл з розширення `model.trt` у `tensorrt`. Невелика кількість обробки кадрів в секунду може бути зумовлений складністю моделі. Моделі YOLO є достатньо важкими, тому для виконання обробки зображення з ідентифікацією об'єктів на Jetson Nano пропонується використовувати спрощені моделі, наприклад такі як YoloV3 tiny, YoloV4 tiny, YoloV5 tiny та схожі [69, 70]. Обробка зображення у роздільній здатності 1920 на 1080 пікселів не гарантує високих показників обробки кадрів в секунду [71]. Проте для забезпечення обробки відео майже 30 кадрів в секунду можна змінити розмір зображення у наборах даних до менших (див. рис. 1.18).

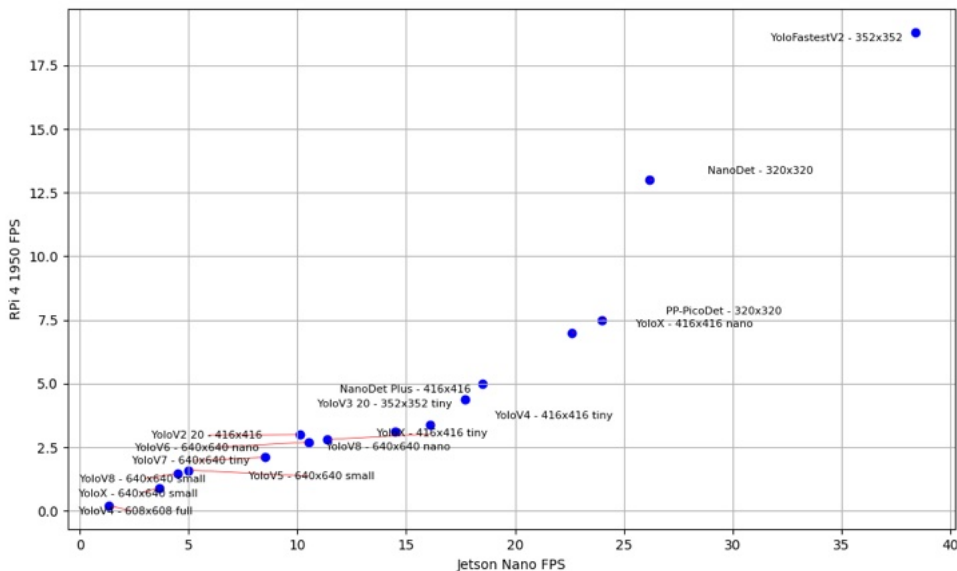


Рис. 1.18. Співвідношення розмірів зображень з моделлю для кількості оброблених кадрів на пристроях Raspberry Pi 4 та Jetson Nano на основі [71]

На основі результатів вимірювань навантажень на графічний процесор під час відповідних обчислень з метою ідентифікації об'єктів на зображенні можна вважати більш ефективніші моделі з меншими за розміром зображеннями, проте точність ідентифікації при цьому є недостатньо високою [72]. А вимірювання ефективності обробки зображень з різною роздільною здатністю показало вплив розміру зображення на результати обчислень кількості кадрів на секунду

використовуючи Nvidia Jetson Nano [73]. Тому всі зображення в наборах даних стандартизовано до середнього розміру 448x448. Обробку фрагментів відео різної довжини виконано з використанням CNN[50], CNN з бібліотекою ONNX Runtime та YOLO на мікрокомп'ютерах Raspberry Pi та Nvidia Jetson Nano. Під час обробки кожного кадру відео на вхідне зображення накладаються маски (червоні виділення) з ідентифікованими дефектами. Обробка зображень і відображення цього процес потребує деякого часу, що спричиняє затримку (Табл. 1.5), що може залежити від складності нейронної мережі та можливостей технічних засобів.

Таблиця 1.5. Час опрацювання фрагменту відео на попередньо натренованих моделях

Мікрокомп'ютер	Тривалість фрагменту відео, секунди	Загальна кількість кадрів	Кількість кадрів в секунду		
			CNN [50]	CNN з ONNX Runtime	YOLO
Jetson Nano	1	32	24.3	26.9	29.4
	9	273	24.6	27.1	29.3
	20	606	24.1	27.5	29.4
Raspberry Pi 4	1	32	22.3	24.2	26.3
	9	273	21.3	23.7	26.1
	20	606	21.4	24.4	25.7

Враховуючи обмежені обчислювальні ресурси мікрокомп'ютерних пристроїв, навчання глибоких нейронних мереж виконувалося у лабораторії технологій штучного інтелекту Львівського національного університету імені Івана Франка та у середовищі Colab [74]. Отриманий модельний файл

(попередньо навчена модель) завантажено на пристрій. Таким чином, проведене розподілене навчання дає змогу пришвидшити час навчання та врахувати обмежені характеристики пристроїв, такі як CPU або GPU, обсяг оперативної пам'яті.

Отож, на затримку обробки можуть впливати великі розміри відео та висока роздільна здатність (високий bitrate), обсяг обчислень нейронної мережі (зменшення енкoderів та декодерів), занадто велика кількість об'єктів для відстеження, наявність артефактів або проблем зі з'єднанням. Проте з оптимізацією нейронної мережі можливі зміни у точності ідентифікації.

У табл. 1.5 такі результати можуть бути зумовлені різними характеристиками апаратної архітектури у Nvidia Jetson Nano (2GB) та Raspberry Pi 4 (4GB), зокрема графічним процесором та оптимізаціями для обробки відео. Nvidia Jetson Nano має графічний процесор NVIDIA Maxwell з підтримкою CUDA, що пришвидшує обробку великого обсягу даних, таких як відео. CUDA дозволяє використовувати паралельні обчислення на графічному процесорі, що особливо корисно для задач машинного навчання та обробки великих обсягів даних. Raspberry Pi 4, хоч і має прийнятні графічні можливості, проте не має такого рівня графічного прискорення та оптимізацій, як у Nvidia Jetson Nano.

## Висновки до розділу 1

У розділі проведено дослідження поточного стану ідентифікації дефектів у ЕІК з урахуванням інтелектуальних та автоматизованих методів, а також задач проектування. Зокрема, однією з частин проведених досліджень є розгляд перспектив застосування сучасних методів машинного та глибокого навчання для досліджуваної тематики. Порівняльний аналіз таких методів з контурними фільтрами показує, що вони за результатами навчання та точністю розпізнавання здатні формувати досить ефективні методики.

Початковим етапом для успішного навчання нейронної мережі є формування збалансованих та великих наборів даних, які у роботі розглянуті для матеріалів типу метали, а також бетонні суміші. Задачі, в яких розглядаються важкодоступні ділянки елемента конструкції з дефектами та автоматизація процесу відбору даних для поверхневого аналізу, розв'язуються з використанням роботехнічних засобів. Описана у відомих наукових працях практика обробки зображень, які містять дефекти у ділянках елементів конструкцій з недостатньою освітленістю, показує ефективність використання додаткових джерел освітлення, інтегрованих у роботехнічні засоби для забезпечення рівномірного освітлення. Також наявність завад може впливати на якість розпізнавання дефектів у інженерних об'єктах і це проявляється відповідно під час обчислення кількості пікселів зображення виділених дефектів. У процесі відповідних перетворень інформації для числового представлення стану дефекту у певному кадрі відеопотоку та під час збору даних з кількох досліджуваних ділянок та їх аналізу прийшли до висновку, що можна сформулювати новий удосконалений комплексний варіант оцінювання ресурсу для системи “металевий підземний трубопровід (МПТ) – СКЗ”.

Окрім аналізу методів обробки зображень поверхневих дефектів у ЕІК проаналізовано також підходи щодо оцінювання їх стану на основі колірних гістограм з урахуванням кількості ідентифікованих пікселів та їх змін у часі.

Відсоткове співвідношення кількості виділених пікселів на зображеннях можуть сигналізувати про критичний стан дефектів на досліджуваній області. Окремий напрямок аналізу зображень для оцінювання зміни станів дефектів пов'язаний з тепловим контролем, який дає змогу отримати групи виділених пікселів і більш точно визначати місця елементів конструкцій з найвищою температурою. Результати дослідження показують, що на основі колірних характеристик поверхневих шарів на зображеннях з урахуванням здійсненої ідентифікації дефектів можна визначати потенційні місця розгерметизації елементів трубопровідних систем.

У роботі порівняно ефективність та точність архітектур НМ CNN та YOLO для різних ЕК та результати використано для аналізу розширеного набору навчальних та тестових даних. На основі метрик для оцінок якості процедури сегментації (IoU, Dice) встановлено, що кращі результати під час ідентифікації дефектів отримані з використанням CNN, що на 35% краще у порівнянні з існуючою реалізацією схожої CNN мережі та на 12% більш ефективно стосовно реалізації з використанням YOLO. З іншого боку, на основі значень обробленої кількості кадрів в секунду встановлено, що більш ефективною є реалізація ідентифікації дефектів з YOLO, якщо використовуємо мікрокомп'ютери Raspberry PI та Nvidia Jetson Nano, зокрема кількісний ефект для них становить відповідно 20.2% та 20.7%. Такого типу результати можуть бути пов'язані з особливостями процедур реалізації методик. CNN має більшу глибину НМ, що дозволяє їй краще розпізнавати складні структури. Натомість більш проста архітектура НМ використовується для реалізації ідентифікації дефектів з YOLO, що дозволяє забезпечити більшу швидкість обробки зображень, але меншу точність сегментації. Крім цього, використані механізми TensorRT і ONNX Runtime дозволяють пришвидшити процедуру ідентифікації дефектів з використанням мікрокомп'ютерів.



Особливості ідентифікації дефектів для корозійних процесів та тріщин у елементах конструкцій також розглянуто у низці відомих наукових робіт і проаналізовано. Ці праці засвідчують, що необхідно наводити зв'язок між результатами діагностування дефектів та методами передачі оброблених зображень.

У дисертаційній роботі для розширення та доповнення розглянутих у першому розділі технологій стосовно задач діагностування ЕІК розроблено ІТ для розпізнавання дефектів у елементах конструкцій на основі реалізованих прогностичних моделей (розділи 2 та 3).

## РОЗДІЛ 2. МЕТОДИ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ЗМІН СТАНІВ ДЕФЕКТІВ У ЕІК

Інтелектуальний аналіз даних передбачає використання статистичних наборів даних для створення моделей на основі яких згодом здійснюватиметься прогноз та прийняття рішення про технічне обслуговування ЕІК.

Одним з напрямків дослідження ЕІК в ІАД є старіння матеріалів. Цей процес є тривалим в часі і ресурсозатратним. Тому для здійснення такого моніторингу змін стану ЕІК необхідно витратити багато часу для формування історичних даних. На основі статистичних даних можна проаналізувати сезонність аварій, найуразливіші місця матеріалів тощо.

Особливої актуальності набула тематика ІАД для ЕІК у час війни унаслідок пошкоджень та руйнування доріг та мостів під впливом надважкої техніки. Станом на липень 2022 року попередньо руйнувань зазнали 23,8 тис. км доріг та 305 мостів в Україні [75]. Крім цього, фактором впливу на зміну стану дефекту у ЕІК є навколишнє середовище. Одним з найбільш поширених видів даних при дослідженні ЕІК є кліматичні (середня температура повітря, вологості) і значення ресурсних ознак (вік матеріалу, наявність ремонтів та швів та інше).

У цьому розділі описано етапи та розробку статистичних моделей для низки ознак дефектоутворення, отриманих із звітів управління з безпеки трубопроводів і небезпечних матеріалів Міністерства транспорту США (PHMSA) [76–78] та досліджень в Україні. Згідно з правилами оператори газорозподільних трубопроводів щорічно подавають звіти про всі небезпечні витоки. У звітах загалом міститься інформація про дату та місцезнаходження, тип механічного кріплення, причина витоку, приблизний рік введення в експлуатацію та інше. Отримані показники використовуються для аналізу довгострокових та короткострокових тенденцій на різних рівнях. В процесі статистичного моделювання розглянуті регресійні моделі з метою формування оцінок зв'язків між основними вищезгаданими параметрами.

Також у дисертаційній роботі запропонований підхід для розв'язування задачі визначення критеріїв найбільшого впливу на зміну стану дефектів у ЕІК та розробленню статистичних моделей. Метою такого аналізу у цій роботі є покрокове проектування системи та відлагодження процесу формування прогностичних моделей. На основі цих результатів спроектовано програмний інтерфейс (Restful API) та здійснено валідацію прогнозу про зміну стану дефектів під час їх ідентифікації на поверхнях камерою готового пристрою (розділ 3).

## **2.1. Аналіз методів та етапів обробки даних про ЕІК**

Загалом, увесь процес обробки та аналізу даних можна відобразити схематично у послідовності кроків взаємопов'язаних між собою (рис. 2.1). Серед використаних засобів для обробки даних є мови програмування Python з бібліотеками Pandas, NumPy, scikit-learn, TensorFlow, PyTorch та R з dplyr, tidyr. Крім цього, ще одним із етапів обробки даних є виявлення та видалення аномалій (викидів) в процесі статистичного моделювання. Наступним етапом попередньої обробки даних є описування кожного із вхідних параметрів створеної конфігурації для подальшого зберігання вихідних даних.

Статистичний аналіз даних виконується у графічному вигляді для візуальної оцінки зміни значень. При цьому основними засобами візуалізації для відображення статистичних залежностей є бібліотека matplotlib та seaborn.

Сформована модель містить категорійні дані певних класів, а саме критичний та не критичний стан дефекту. Переведення в числові формати здійснюється за допомогою процесу кодування. Кодування категоріальних ознак може бути виконано з використанням метод фіктивних змінних (one-hot encoding), кодування міток label encoding, чи інших методів. Метод фіктивних змінних полягає у зміні текстових даних які позначаються якісь категорії (наприклад, дефект чи не дефект та інше) на бінарні змінні. Використання методів для балансування (вирівнювання) класів, якщо є дисбаланс. Серед

великої кількості параметрів виділяються всі ті, які будуть найбільш важливі з точки зору кореляції значень між собою для подальшого аналізу. Використання методів відбору ознак (Feature Engineering) для визначення важливих ознак на основі кореляційної матриці. Цей процес необхідний для видалення зайвих ознак для зменшення їх кількості. (рис. 2.1).

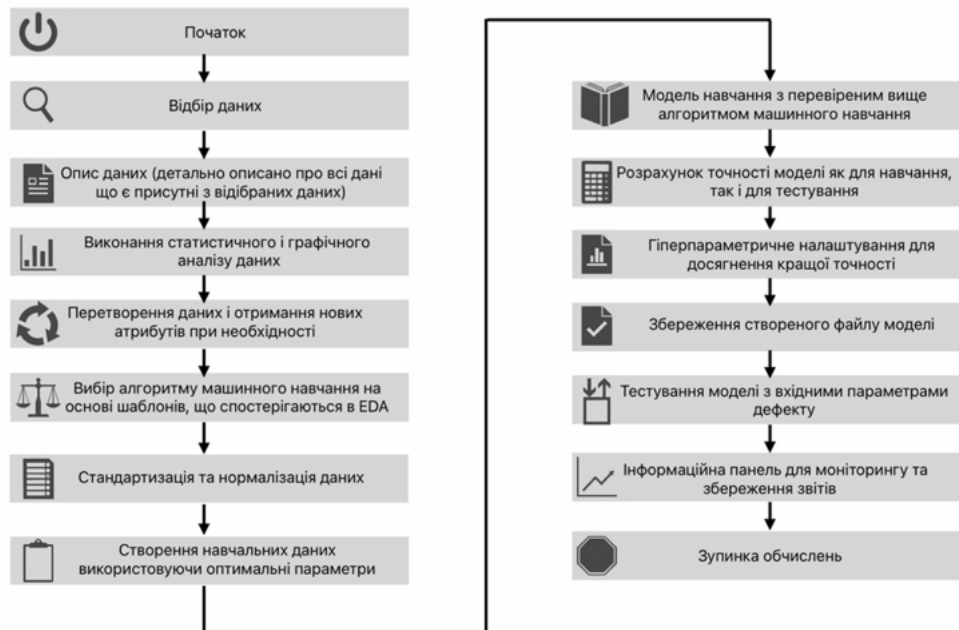


Рис. 2.1. Покрокова схема опрацювання даних та аналізу дефектів

На основі шаблонів даних можна вибрати алгоритм машинного навчання у розвідковому аналізі даних (EDA – Exploratory Data Analysis). Такий аналіз дозволяє на ранніх етапах обробки даних визначити загальні тенденції або властивості даних. Часто дані потребують перетворення до стандарту та нормалізації даних для приведення їх до єдиного формату та структури. Таке формування наборів даних для навчання потребує підбору оптимальних параметрів таких як: глибина дерева (*max\_depth*), максимальна кількість ознак для розгляду під час розгалуження (*max\_features*), мінімальна кількість вибірок для розгалуження (*min\_samples\_split*) та інше.

Для вибору алгоритму машинного навчання варто порівняти різні реалізації на одному незмінному наборі даних. З недостатньою кількістю даних

загальний набір може ділитися на два (навчальний та тестовий) у заданій відсотковій пропорції. Результати отриманої точності навчених моделей перевіряються відразу і дозволяють оцінити ефективність використаних алгоритмів. Також порівняння точності після виконання навчання дає змогу підібрати оптимальні гіперпараметри моделі. Для багаторазового використання даних та результатів навчання, усю інформацію зберігають для подальшого порівняння або візуалізації.

Валідацію результатів навченої моделі можна здійснити вже на основі реальних даних, визначивши зміни стану дефекту.

## 2.2. Опис та формалізація інформативних показників ЕІК

З відкритих джерел адміністрації з безпеки трубопроводів і небезпечних матеріалів Міністерства транспорту США можна проаналізувати статистичні дані аварійних випадків витoku рідин [76], розподілу газу [77], несправності технічного обладнання з 2011 року до 2021 (131553 записи) [78].

Причини виникнення дефектів можна оцінити на основі історичних даних із звітів аварій про механічні пошкодження об'єктів [78]. Найчастіше, як видно з рис.2.2, причиною аварій було обладнання, матеріали та природні сили.

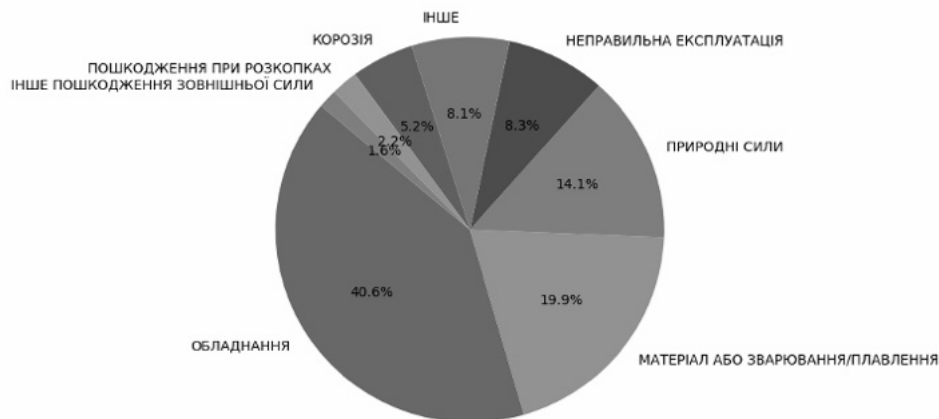


Рис. 2.2. Кругова діаграма відсоткового співвідношення причин аварій.

*Примітка: побудував дисертант на основі <https://www.phmsa.dot.gov>*

Як відомо з роботи [79], під час експлуатації виникає найбільша кількість дефектів. У праці [80] в мостових конструкціях основними матеріалами у пошкоджених конструкціях є метал та цемент як показано на Рис. 2.2. Відповідно у різних матеріалах виникають дефекти схожі за формою на тріщини чи пітингова корозія у металах. У будь-якому разі, це спричинює руйнування чи зупинки роботи системи чи об'єкта. Для моніторингу змін у матеріалах використовують ІАД, які дозволяють оцінити ризики на основі аналітики даних та сформувати певні залежності для уникнення таких ситуацій у майбутньому.

У праці [81] стверджують, що найчастіше причинами аварії в Україні є: фізичні дії (незаконні врізання), відхилення під час ремонту та будівництва, корозійні пошкодження, якість труб і устаткування, помилкові дії персоналу.

Згідно до наказу №36 Державного комітету України по водному господарству про затвердження Методики визначення корозійного стану трубопроводів зрошення та водопостачання технічна документація повинна включати підбір проектної та виконавчої документації, даних експлуатуючих організацій, гідрогеолого-меліоративних експедицій, відомості про матеріали попередніх обстежень та про заходи стосовно захисту від корозії. Тобто дані вологість ґрунтів, рівень промерзання земного покриву, тип матеріалу та дата будівництва ЕІК безпосередньо впливають на довговічність функціонування таких типів систем в цілому. Дані про кліматичні показники можна отримати з погодних сервісів, зокрема використовуючи відкриті бази даних [82].

Питання формалізації параметрів як набір поверхневих параметрів у моделях розглядалися у роботах [83, 84]. При цьому основними параметрами для кореляційно-регресійного аналізу обрано глибину дефекту та температуру.

У праці [85, 86] розглянуто огляд причин пошкодження труб та факторів, що впливають на довговічність їх експлуатації. Таким чином, отримані результати у вищезгаданій роботі дають змогу стверджувати про важливість таких факторів як: температура, ступінь корозії та інші фізичні показники.

Оскільки додатково зібрані параметри створеної моделі є важливі при роботі формуванні наборів даних. Тоді розширена кількість параметрів, таких як форма труби, товщина труби, внутрішня або зовнішня поверхня, наявні зварні шви, з'єднання та попередній ремонт представлені в праці [87] дозволить більш точно здійснити опис моделі.

Однак цей набір параметрів можна доповнити деякими параметрами з роботи [88], такими як: вік, погода, захист, тиск, гідрогеологія, особливості ґрунту. Ці параметри необхідно враховувати під час формування статистичної моделі. Особливості експлуатації та параметри, що впливають на міцність металу та системи контролю описані в праці [10]. Методика оцінки довговічності матеріалів описана в праці [89]. Ці параметри і фактори стосуються в основному підземних трубопроводів, які необхідно враховувати під час створення статистичних таблиць.

Деякі особливості експлуатації та оцінка старіння за роками з урахуванням сезонних температур описані в роботі [12]. Після аналізу робіт та параметрів сформовано Табл. 2.1, яка містить деякі параметри важливі для статистичної моделі оцінювання стану поверхневих дефектів.

Таблиця 2.1. Важливі параметри для статистичної моделі оцінки стану дефекту

Параметр	Опис	Значення	Співставлення з полями у наборах даних
Назва матеріалу	Основний матеріал елемента ЕІК	Бетон, метал, бетон з арматурою	FITTING_MATERIAL_TEXT
Тип матеріалу	Суміш або назва конкретного типу враховуючи особливості матеріалу	Метал: сталь, карбон Бетон: нановапняк, нанокремнезем,	MODEL_NUMBER

		наноглина	
Розмір	Діаметр чи довжина об'єкта	Від 400 до 900 мм	FIRST_PIPE_NOMINAL_SIZE_TEXT, SECOND_PIPE_NOMINAL_SIZE_TEXT
Місцерозташування об'єкта	Найближче місто до аварійної ділянки	Середні значення погодніх особливостей	HQ_ADDRESS_CITY
Вік	Тривалість експлуатації конструкції	кількість років (наприклад, 30)	YEAR_INSTALLED
Розташування (для труб)	Логічне значення щодо підземного або надземного використання	підземне/наземне	LEAK_LOCATION_A_TEXT
Захист	Наявність антикорозійного захисту металу або додаткового покриття бетону	присутній/відсутній	1 - присутнє 0 - відсутнє
Вологість	Наявність вологи з навколишнього середовища (наявність водойм, болотистість)	висока, середня, низька	LEAK_LOCATION_B_TEXT
Можливе розташування дефекту	Зазначення місця огляду або наявності дефектів у минулому	внутрішній/зовнішній	LEAK_OCCURRED_TEXT
Можливий тип	Найбільш часті	Витік, корозія	LEAK_CAUSE_T



дефекту	дефекти на основі історичних даних під час ремонтів	або тріщина	ЕХТ
Рівень ризику	Рівень ризику, який може бути з урахуванням зазначених параметрів	Високий, середній або низький	2 - високий 1 - середній 0 - низький

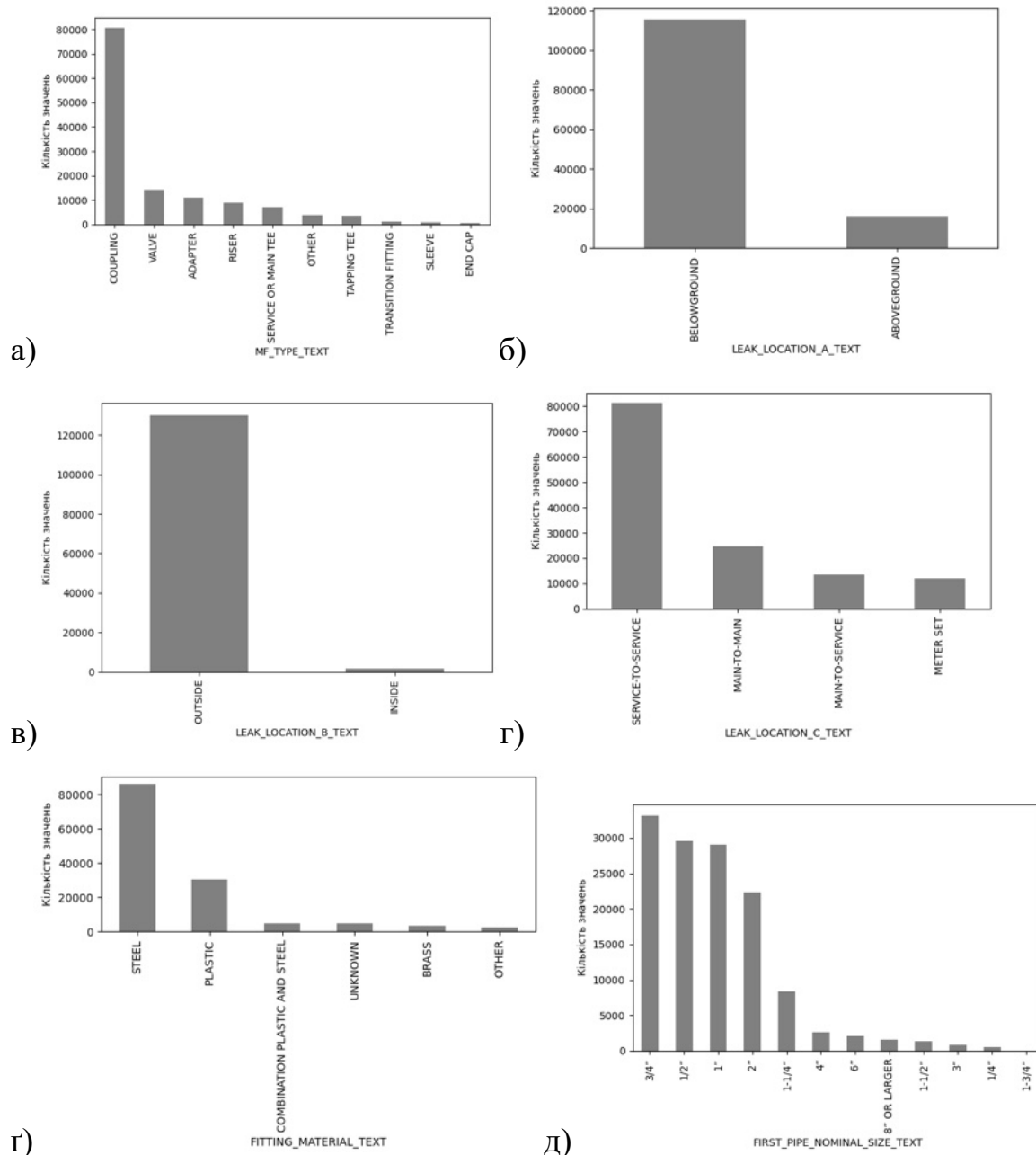
На основі порівняльного аналізу праці [81-88] зроблені висновки, які підтверджують актуальність роботи і вказують на необхідності розробки нових та вдосконалення існуючих моделей, методів і інформаційних технологій для проведення наукових і експериментальних досліджень, пов'язаних з обробкою та аналізом послідовності даних для встановлення найбільш вразливих місць елементів конструкцій та оцінки стану дефекту в задачах технічної діагностики конструкційних матеріалів.

Для формування систем прийняття рішень, які дозволяють здійснювати прогноз зміни стану дефектів необхідно враховувати додатково ризику під час впровадження ІТ та здійснювати управління проектами в ІТ [89, 90].

### **2.3. Статистичний аналіз інформативних параметрів**

Проведення статистичного аналіз даних дозволяє отримати важливу інформацію щодо зміни стану дефектів, яка допоможе зрозуміти та вдосконалити якість оцінки ресурсу об'єкта та безпеку систем. Часто аварійність зумовлена обмеженим ресурсом ЕІК, що впливає на функціонування системи в цілому. Для відстеження змін, дані з статистичного аналізу можна використати для частотного (рис. 2.3), трендового, кореляційного, географічного аналізу даних про дефекти [78], їх класифікацію та отримувати інші залежності та показники. Отримані результати на основі описаних даних можуть бути корисними для широкого спектру завдань: виявлення змін в часі при різній кількості дефектів,

побудови розподілу дефектів за типами та категоріями, визначення залежностей між дефектами та іншими факторами (на основі регресійного аналізу для оцінки впливу різних факторів), вивчення розташування дефектів на карті для ідентифікації географічних зон з підвищеним ризиком та інше. Статистичний аналіз дефектів може бути корисним для виявлення та вирішення проблем, визначення важливості ремонтів ЕІК, а також для вдосконалення процесів розробки чи виробництва (рис. 2.3).



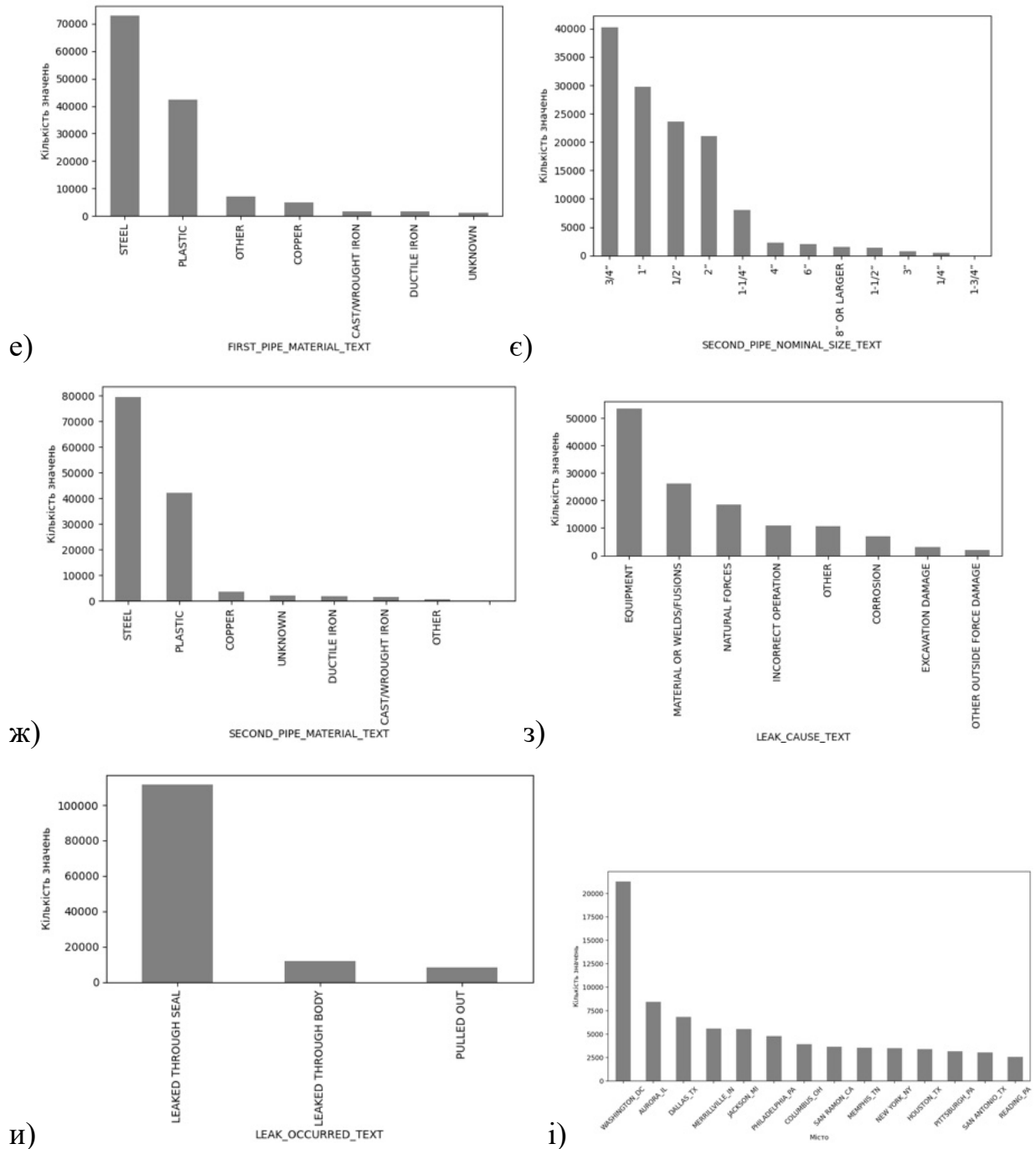


Рис. 2.3. Частотний аналіз значень параметри елементів конструкцій: а) типу, б) розташування об'єкта, в) розташування дефекту, г) тип з'єднання, ґ) тип матеріалу з'єднання, д) номінальний розмір першої частини пошкодженого елемента конструкцій, е) тип матеріалу першої частини пошкодженого елемента конструкцій, є) номінальний розмір другої частини пошкодженого елемента

конструкцій, ж) тип матеріалу першої частини пошкодженого елемента конструкцій, з) причина пошкодження, и) тип пошкодження, і) міста

*Примітка: побудував дисертант на основі <https://www.phmsa.dot.gov>*

Визначення і обробка пропущених значень в аналізі даних є важливим етапом, який потребує ретельно вибраних кроків для прийняття наступних рішень про важливість параметрів моделі. Це зумовлено тим, що ці дії можуть суттєво вплинути на результати та висновки. Пропущені значення можуть призвести до втрати важливої інформації та зниження точності аналізу. Також, без обробки пропущених значень може бути складно побудувати моделі або виконати статистичний аналіз, оскільки багато алгоритмів та методів вимагають повного набору даних [80]. Варто також врахувати забезпечення консистентності та репрезентативності вибірки, що може призвести до неналежного визначення особливостей або закономірностей в даних.

Проведення аналізу даних може бути некоретним також при відсутності значень у даних, що спричиняє унеможливлення встановлення залежностей між даними і можуть впливати на результати та висновки.

Тому найкращою стратегією уникнення незаповнених значень є заміна недійсних записів найбільш частими значеннями в цій функції. Алгоритми машинного навчання не можуть працювати з рядками або символічними даними. Тому потрібно буде змінити ці значення на числові, перш ніж передати набір даних в алгоритм МН для цього використовують кодування міток. Під час цього процесу відбувається присвоєння кожному класу даних унікальний номер, починаючи з 0. Результатом таких дій є отримання двох або більше наборів даних. Набір даних поділяються на навчальний та тестовий. Після цього відбувається підбір найефективнішого відповідно до точності класифікатора моделі з існуючих реалізацій у бібліотеці Scikit.

Як показано на частотних графіках (рис.2.3) деякі мають значення невідоме (unknown) або можуть бути пустими. На рис. 2.4 показано приклад наборів даних з деякими пропущеними значеннями.

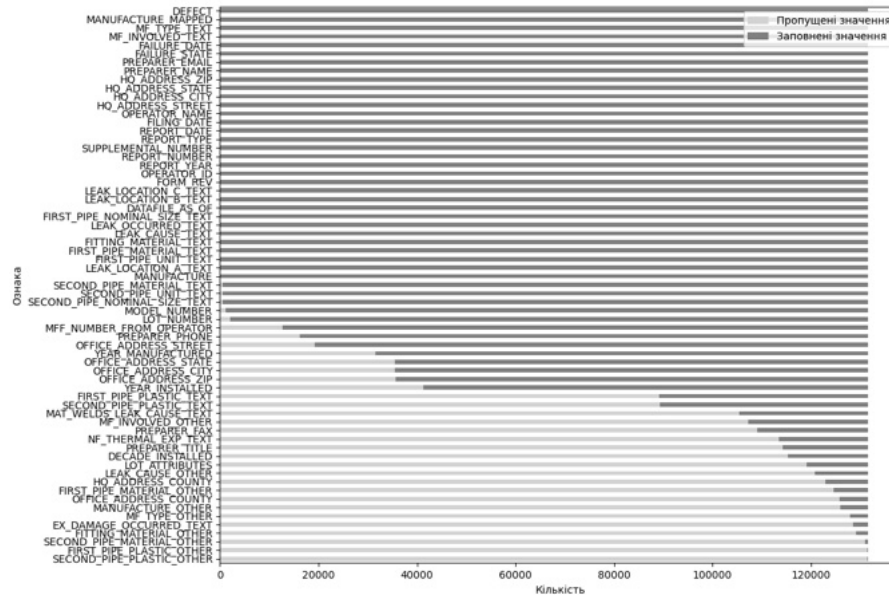


Рис. 2.4. Кількість пропущених та заповнених даних для набору даних

## 2.4. Аналіз залежностей між інформативними даними

На основі історичних даних про приблизні роки виробництва трубопроводів та кількості аварій зображено тенденції задокументованої кількості аварій по десятиліттях з кінця 1880-х до початку 2020-х років. Тому спостерігається збільшення кількості ремонтів до другої половини XX століття, яке можливо пов'язане із зношенням старіших трубопроводів. Вік матеріалу може негативно впливати на зміну стану дефектів і, відповідно, аварійних ситуацій при функціонуванні досліджуваних об'єктів. Такі тенденції можна візуально оцінити за допомогою побудови стовпчастої діаграми кількості аварій на трубах за період з 1880 по 2020 рр. в залежності від тривалості часових інтервалів експлуатації (рис. 2.5).

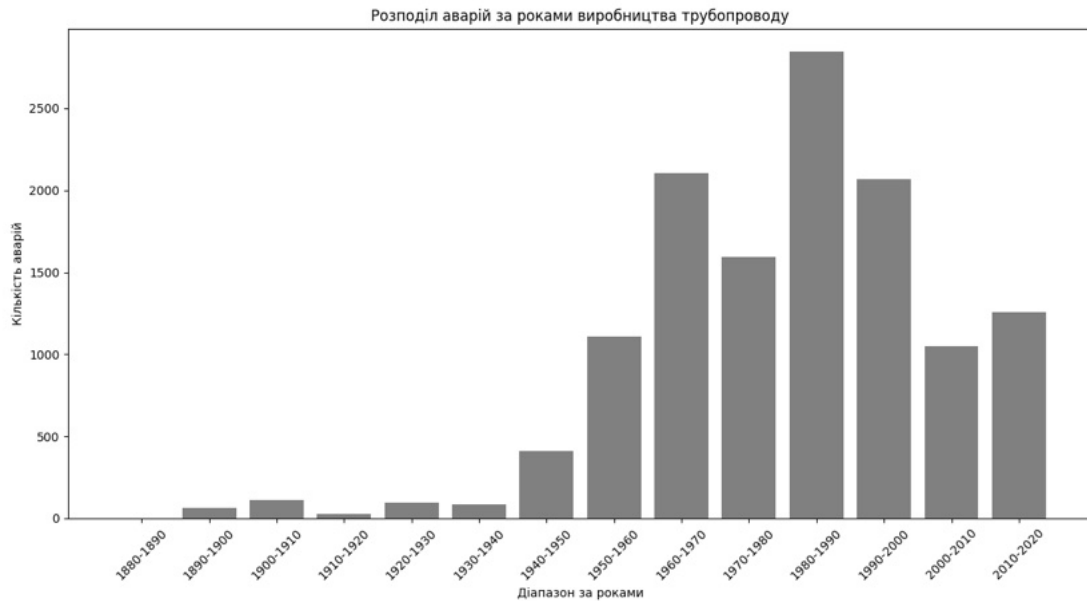


Рис. 2.5. Розподіл кількості аварій за роками виробництва трубопроводів

*Примітка: побудував дисертант на основі <https://www.phmsa.dot.gov>*

Схожим способом можна проаналізувати вплив зовнішній сезонних впливів на функціонування об'єкта. На рис. 2.7 показано, що у зазвичай більш теплі літні місяці кількість аварій менша у приблизно половину від зимових. Таке представлення інформації у графічному вигляді дає змогу проаналізувати тенденції та переконатися у важливості вибраних кліматичних параметрів моделі. Як показано у роботі [12] на прикладі з'єднання нанобетону з арматурою, що вплив сезонних температур дає змогу оцінити зміну стану дефекту. Аналіз критеріїв якості та їх взаємодію приведено у процесі математичного моделювання (розділ 4).

Згідно з даними [81], яке базується на відкритих даних про функціонування нафтопроводів України про введенних у експлуатацію з початку 1960-ті до початку 1970-тих років відносяться більше 30 %, а в наступне десятиліття близько 50 %. Решта даних щодо років введення в експлуатацію показують приналежність до новіших часових проміжків і ці об'єкти можуть бути менш схильні до виникнення дефектів і відповідно аварійних ситуацій.

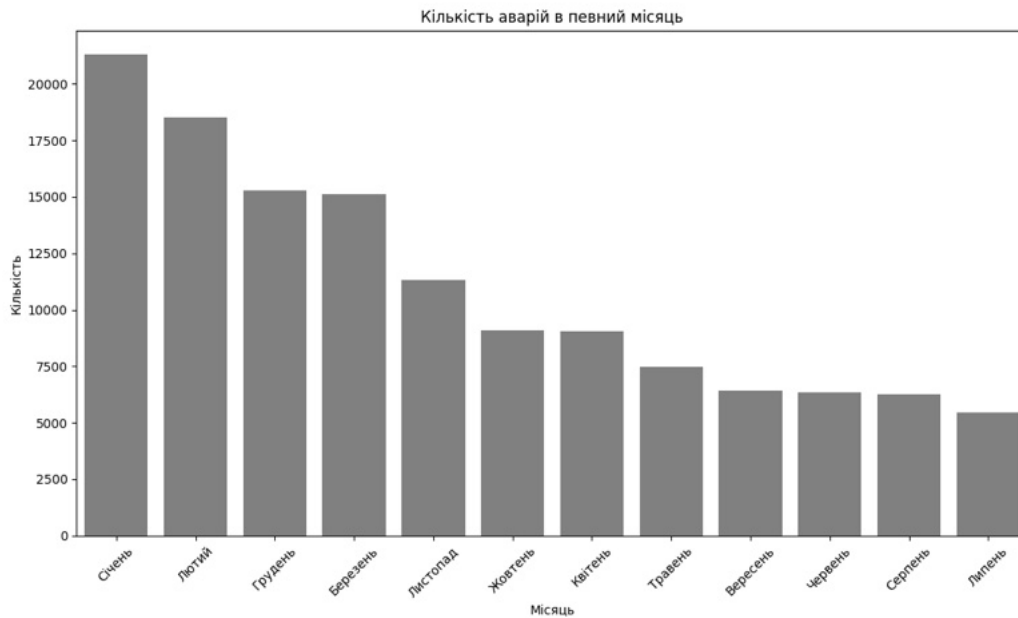


Рис. 2.6. Розподіл кількості аварій за місяцями року

*Примітка: побудував дисертант на основі <https://www.phmsa.dot.gov>*

Візуалізація даних про залежність частоти аварій до місяця року дає змогу стверджувати важливість врахування в параметрах поточну дату дослідження. Оскільки температури навколишнього середовища впливає і пришвидшує руйнування матеріалу, то під час прогнозування варто звернути увагу на цей параметр також.

Таким чином, такі розподіли (Рис 2.5 і Рис. 2.6) можна використати для статистичного моделювання та встановлення залежностей між можливими параметрами. Проте, більш складнішим в реалізації є процес прогнозування зміни стану дефектів.

Попередній аналіз зібраних даних можна виконати за допомогою розвідкового аналізу даних (РАД). Деякі параметри можна виміряти безпосередньо в досліджуваному об'єкті, наприклад, рівень вологості. Серед інформативних параметрів заздалегідь визначені такі як вік матеріалу, кількість ремонтів тощо. Коробкова діаграма (рис. 2.7) дозволяє оцінити рівень пошкодження для різних видів матеріалів. Відповідні показники дозволяють оцінити наявність найбільш вразливих місць і перевірити результати прогнозу на

основі запропонованого в даній роботі методу розпізнавання дефектів в рамках обмежених розрахунків.

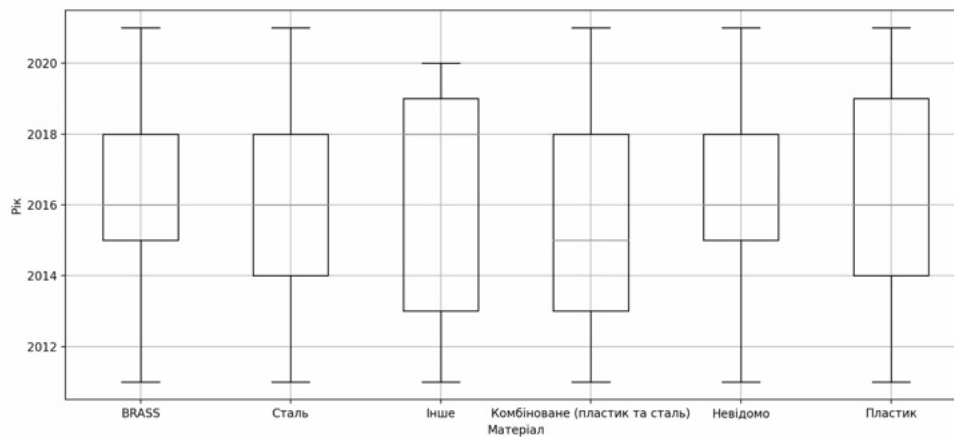


Рис. 2.7. Діаграми розмаху пошкоджень за видами матеріалів

З проведеного аналізу факторів впливу на функціонування ЕІК варто звернути увагу на кліматичні особливості досліджуваної ділянки (зокрема, показники вологості або температури). Зазвичай, відкриті сервіси [82] містять великі набори даних, що дозволяють підсумувати збережену інформацію про погоду та отримати одразу середньостатистичні значення відповідних показники за весь період досліджень. Погодні особливості у межах географічного розташування досліджуваного об'єкта визначені на основі середньорічних показників вологості та температури, які впливають на зміну станів дефектів на поверхнях ЕІК (Табл. 2.2).

Таблиця 2.2. Визначення середньої річної відносної вологості та температури у населеного пункту із зафіксованими аварійними ситуаціями

Місто	Кількість записів	Середня річна відносна вологість, %	Середня річні температура, °F (°C)
WASHINGTON (DC)	21271	64	44.1 (24.5)
AURORA (IL)	8447	69	51.0 (28.3)



DALLAS (TX)	6824	59	39.7 (22.1)
MERRILLVILLE (IN)	5610	70	49.1 (27.3)
JACKSON (MI)	5558	68	35.3 (19.6)
...			
CHELSEA (OK)	1	61	45.0 (25.0)
WAYLAND (IA)	1	66	53.0 (29.4)
MORTON (IL)	1	67	50.1 (27.8)

Серед аварійних ділянок поблизу міст знайдено майже 300 унікальних записів найближчих населених пунктів та додано значення середньої річної відносної вологості та температури.

Модель стану дефектів будується на основі аналізу даних, які містять інформацію про стан елементів або конструкцій. Ця модель використовується для прогнозування або виявлення потенційних дефектів або виробничих проблем на основі оцінки ризиків імовірності їх виникнення.

Таким чином, сформована Табл. 2.3 з заданими умовами вищенаведеного аналізу даних на основі кількох визначених категорій параметрами впливу на зміну стану дефектів.

Таблиця 2.3. Категорії статистичних параметрів залежно від стану дефектів

Параметри		Стан дефекту	
		Критичний	Не критичний
Рік	до 1960		+
	1960-нині	+	

Вологість	до 65%		+
	вище 65 %	+	
Температура	до 43.8 °F		+
	вище 43.8 °F	+	
Матеріал	Сталь	+	
	Пластик		+
	Інше		+
Місяці	зимові	+	
	осінні	+	
	весняні	+	
	літні		+

Ймовірність аварії може суттєво залежати від поточного стану дефекту, де мірою цього стану може бути визначений вплив на безпеку та надійність системи. Наприклад, якщо встановлений дефект має невеликі розміри та низьку ймовірність розширення, його вплив може бути оцінений на рівні 5, що вказує на мінімальний ризик. У той час як критичний стан дефекту із значними розмірами та високою ймовірністю подальшого розростання може мати максимальний вплив (рівень 1) і збільшувати ймовірність аварії.

Оцінка впливу від 1 до 5 може використовуватися для визначення необхідності та пріоритетності вжиття заходів щодо ремонту, заміни або обслуговування об'єкта. Відповідно, чим вищий рівень впливу, тим більш термінові та значущі можуть бути дії для забезпечення безпеки та уникнення аварій як показано у Табл 2.4.

Таблиця 2.4. Категорії ризику залежно від критичного стану дефектів та ймовірності аварій

Стан дефекту	Ймовірність аварії		
	Висока	Середня	Мала
Критичний	Дуже високий ризик (1)	Високий ризик (2)	Середній ризик (3)
Не критичний	Середній ризик (3)	Малий ризик (4)	Дуже малий ризик (5)

Для кількісного визначення впливу можна використовувати регресійний аналіз, що дозволяє оцінити вагомість внеску кожного параметра у змінні результати. Наприклад, якщо аналізуються параметри, такі як рік, вологість, тип матеріалу та місяць у контексті ефективності трубопроводу, можна використовувати статистичні методи для визначення та як ці фактори взаємодіють і які з них найбільше впливають на ефективність. Це допоможе приймати обґрунтовані рішення щодо управління та оптимізації трубопровідної системи з урахуванням мінливих умов вологості, типу матеріалу та інших факторів протягом року.

## 2.5. Регресійний аналіз важливості ознак моделі ресурсу ЕІК

Серед набору даних можна виділити деякі менш важливі дані для багатокритеріального аналізу даних у процесі кореляційного порівняння. Цей процес спрямований на забезпечення якості та адекватності даних для подальшого використання. Основні етапи процесу очищення та підготовки даних включають: збір даних з різних джерел (таких як бази даних, файли, API тощо), обробка пропущених значень, виправлення помилок у даних (неточності, аномалії, викиди), нормалізація та стандартизація, генерація ознак (feature engineering), збалансування вибірки для уникнення перекошу, розділення на

тренувальний (80% - 105242 записи) та тестовий (20% - 26311 записи) набори, валідація та перевірка якості.

Огляд робіт в останніх періодичних виданнях дає змогу стверджувати, що дедалі більше реалізацій прогностичного аналізу про зміни стану дефектів та визначення важливості параметрів розглядається через призму використання машинного навчання і аналізу інженерії ознак. До того ж можна вважати, враховуючи різноманіття праць, що залежно від предмета досліджень наукові праці можна умовно поділити на три групи. Перша – статті, які присвячені визначенню найважливіших параметрів моделей. У цьому напрямі доцільно виділити праці [85-88].

Другою групою є дослідження спрямоване на застосування нечіткої логіки у таких працях [91-95].

Третьою є праці [96-100] здійснення безпосередньо прогнозування на основі технологій машинного навчання.

Попри те, що існує велика кількість наукових змістовних праць, у яких розглянуто проблеми вибору параметрів для моделей різної тематики, є підстави стверджувати, що відсутні універсальні способи для класифікації з метою прогнозування зміни стану дефектів в контексті аналізу даних. Отже, дослідження у цьому напрямку можна вважати досить фрагментарними та несистемними. Здебільшого для здійснення прогнозування кожної предметної області виконують відбір даних та опрацьовують їх відповідно до специфіки.

Результати проведеного статистичного аналізу з урахуванням стану матеріалів свідчать про актуальність досліджень у цьому напрямку. Ці історичні дані є корисними в контексті формування наборів статистичних даних [22]. Покращити якість даних та підвищити ефективність використаних моделей машинного навчання можна з використанням інженерії ознак (Feature Engineering) [83, 101, 102]. Перевагою застосування якої є зниження обсягу даних з виділенням найбільш важливих, що у деяких випадках спрощує модель і

пришвидшує час навчання. Також дає змогу запобігти перенавчанню з видаленням зайвих параметрів (шумів). У процесі підготовки наборів даних відбувається перевірка наявності повторів, заміни текстових значень на цифрові для використання класифікаторів машинного навчання [100]. Навчальний набір даних включає позитивний (критичний стан дефектів) і негативний сценарії (не критичний стан дефектів).

Загалом існують різні моделі машинного навчання, які використовуються для різних видів задач. Вибір оптимального підходу здійснений використовуючи порівняльний аналіз показників точності наступних моделей:

- Linear Regression (лінійна регресія) дозволяє здійснюється пошук лінійної залежності між вхідними факторами і вихідною змінною.
- Logistic Regression (логістична регресія) використовує логістичну функцію для оцінки ймовірності приналежності до певного класу.
- Decision Tree (дерево рішень) приймає рішення, розглядаючи низку факторів і просуваючись вглиб, враховуючи умови розділення.
- Random Forest (випадковий ліс) є схожим до дерева рішень, проте це ансамбль багатьох дерев рішень, щоб зменшити перенавчання та підвищити точність.
- XGBoost (eXtreme Gradient Boosting) це ансамбльний метод градієнтного підсилення, який поєднує багато слабких моделей для створення потужної ансамблевої моделі.

Різниця між цими моделями полягає в тому, як вони взаємодіють з даними та вирішують конкретні завдання. Лінійна регресія підходить для простих лінійних залежностей, тоді як логістична регресія - для класифікації. Дерева рішень, випадковий ліс і XGBoost використовуються для більш складних випадків, коли залежності можуть бути нелінійними чи взаємозалежними. Ансамблеві методи, такі як Random Forest і XGBoost, часто дозволяють покращити точність моделі за рахунок комбінації декількох "слабких" моделей.

Аналогічно як у [23] можна зобразити ступінь важливості врахованих параметрів у моделі (рис. 2.8).

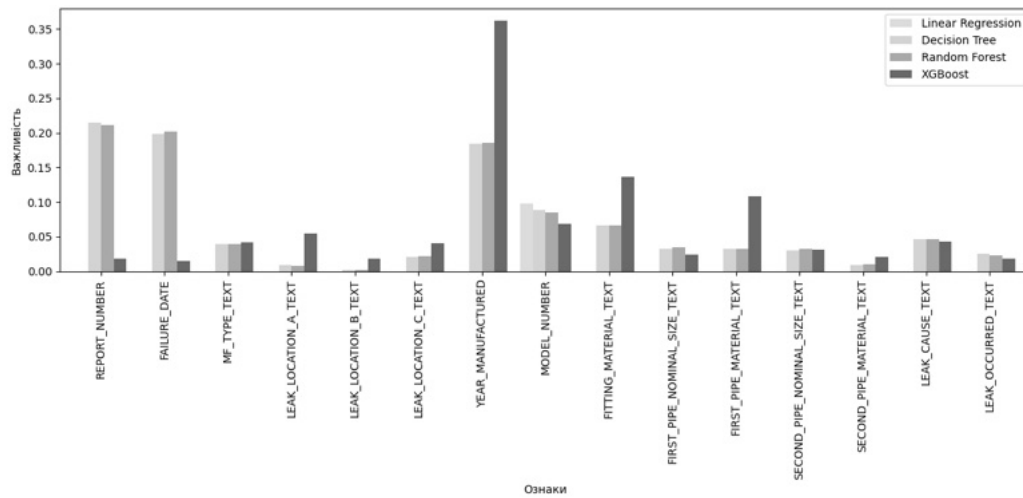


Рис. 2.8. Порівняння важливості ознак між моделями на даних з [78]

Дерева рішень є простими та інтуїтивно зрозумілими моделями, однак вони є моделями з високою дисперсією, тобто незначна зміна даних про групу може призвести до низької продуктивності під час тестування, оскільки вони намагаються знайти складні зв'язки в даних. Незважаючи на те, що модель має стабільні результати щодо даних підготовки та тестування, ця модель має низьку чутливість і високу специфічність.

У контексті машинного навчання, дерева рішень використовуються для класифікації або регресії, і вони можуть стати доволі громіздкими та складними. Візуалізація дерев рішень може допомогти зрозуміти їхню структуру та процес прийняття рішень. Серед переваг візуалізації дерев рішень є читабельність з показаними функціями та умовами для поділу та вибору того чи іншого шляху. Крім цього, на основі глибини гілок можна оцінити перенавчання в дереві рішень, які можуть бути пов'язані з недосконалістю підібраних умов. Також у процесі відповідного аналізу інформації можна ідентифікувати кожну гілку і переглянути параметри, які мають ключове значення для прийняття рішень або найбільш важливі гілки для вхідних даних [103]. Отож, ці параметри оцінені за допомогою

програми написаної мовою програмування Python, бібліотеки scikit-learn для навчання моделі та graphviz для візуалізації графа (рис. 2.9).

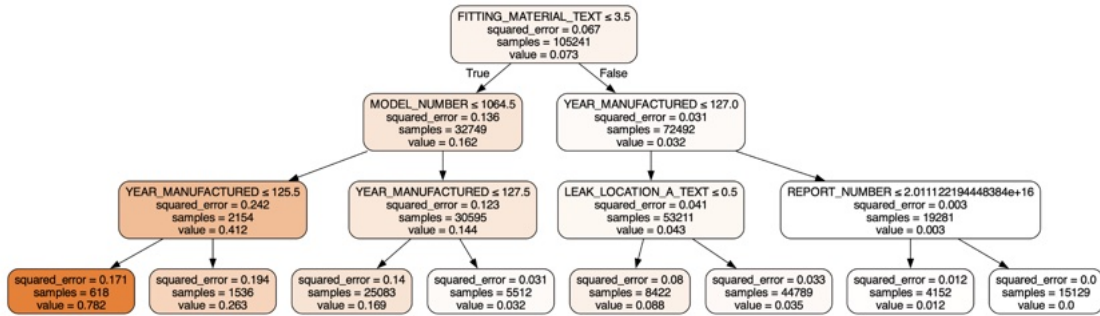


Рис. 2.9. Візуалізації дерева рішень для заданої глибини 3

Оптимальне значення глибини дерева отримано ітеративним шляхом пошуку найменших похибок зі збільшенням глибини як показано на Рис. 2.10.

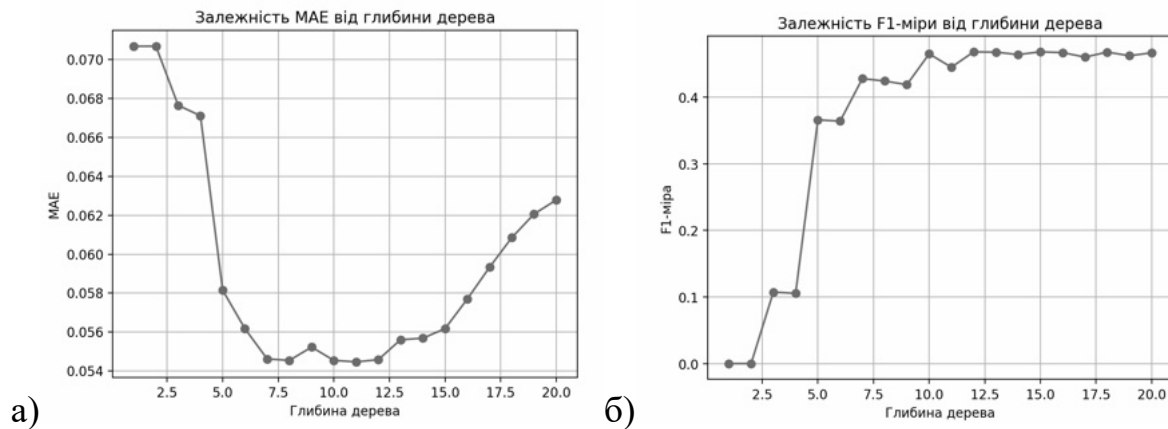


Рис. 2.10. Графіки залежності глибини дерева моделі у класифікаторі від:

а) середньої абсолютної похибки, б) F1-міри

За замовчуванням дерево розширюється до того часу, поки кожен лист не буде містити менше числа зразків, ніж заданий параметр мінімальної кількості зразків для розгалуження на дочірні вузли, або до того часу, поки всі листи не будуть чистими (міститимуть об'єкти лише з одного класу). Пошук оптимального значення глибини дерева є важливою частиною налаштування моделі дерева рішень в машинному навчанні. Визначення оптимальної глибини дерева дозволяє досягти балансу між недонавчанням і перенавчанням. Оптимізація глибини

дерева часто проводиться експериментально, за допомогою методів крос-валідації або пошуку по сітці (grid search). Це дозволяє визначити той рівень глибини, який найкраще підходить для конкретного набору даних та задачі. Згідно з Grid Search найкращі результати отримані з параметром швидкості навчання (learning rate) алгоритму градієнтного бустінгу XGBoost 0.1 та визначено кількість дерев, які будуть побудовані під час навчання моделі. 300 з оцінкою 0.36. "Mean Test Score" (середня оцінка тестового балу) використовується для оцінки продуктивності моделі під час к-згинаючої перехресної перевірки. Вона представляє собою середню точність (або іншу метрику продуктивності) моделі на тестових наборах даних для кожного згину (фолду) під час виконання к-згинаючої перехресної перевірки (рис. 2.11).

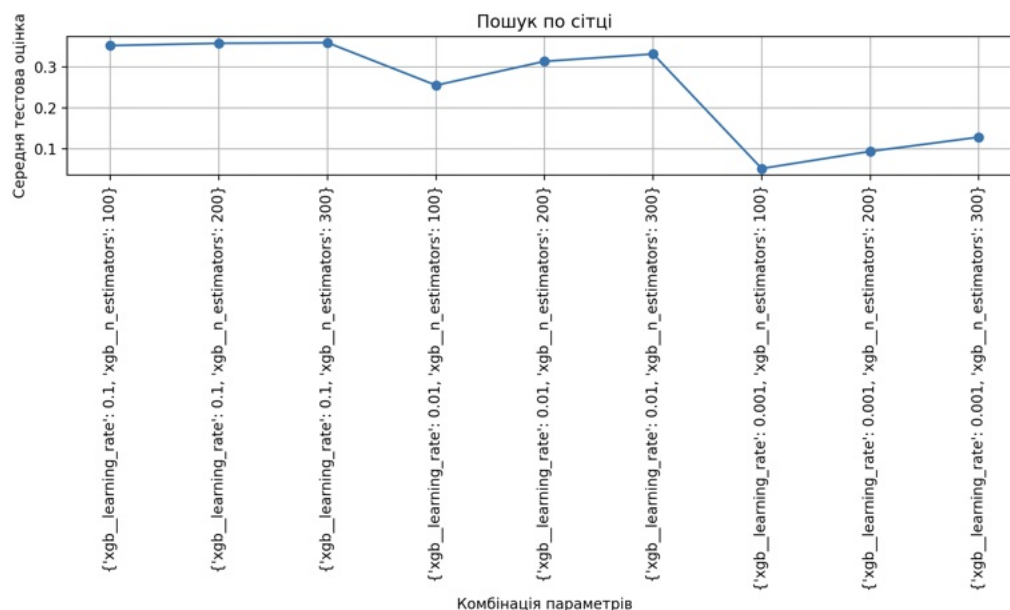


Рис. 2.11. Методи покращення градієнтного підсилення за допомогою к-згинаючої перехресної перевірки

Загалом проаналізована вибірка показників вказує на те, що загалом з позиції цінностей раціонально обрані параметри [102, 103]. Оптимальні параметри, отримані в результаті інженерії ознак, можуть сприяти покращенню ефективності та точності моделі.



Розділення основних параметрів моделей для трубопроводів, мостів та доріг є важливим з огляду на те, що ці інженерні структури мають свої унікальні характеристики та вимоги до проектування та управління. Кожен тип інженерної структури вимагає специфічних підходів до аналізу, проектування та управління, і тому має свої власні ключові параметри, як було вище згадано.

## 2.6. Формалізація станів дефектів на поверхнях

Прогностичні моделі допомагають вирішити завдання прогнозування та аналізу тенденцій у задачах класифікації. Розставлення міток в прогностичних моделях є важливим етапом, де мітки вказують на цільові значення, які модель намагається передбачити.

Розставлення міток у даних використовується після процесу очищення та підготовки даних, статистичного моделювання та проміжної візуалізації [104]. Відповідно до сформованих статистичних категорій (Табл. 2.2) можна класифікувати дані відповідно до цих співвідношень.

Крім того, використання бінарних чисел (0 і 1) дозволяє більш ефективно та стандартизовано представляти та обробляти інформацію, зокрема стосовно двох станів ("не критичний" та "критичний" відповідно). У задачах класифікації, модель може навчатися на цих бінарних значеннях для класифікації об'єктів на два класи.

Діаграма розмаху (Рис. 2.12) кількості даних для критичного і не критичного станів дефекту відповідно до умов заданих у Табл. 2.3.

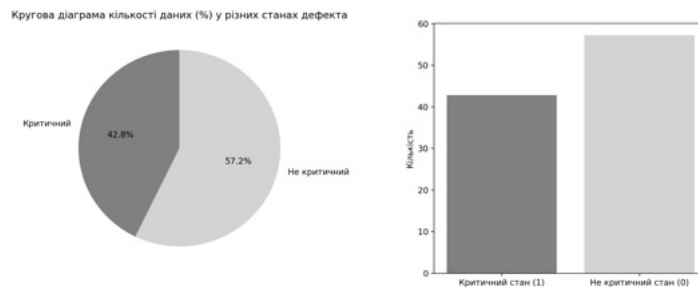


Рис. 2.12. Кількість даних із значенням критичного і не критичного значень станів дефекту

Набори даних є достатньо збалансованими, після цього виконується навчання моделей та оцінювання точності.

## 2.7. Методи оцінювання точності прогнозу змін станів поверхневих дефектів

Одним із напрямків інтелектуального аналізу даних є прогнозування. Прогноз може здійснюватися з використанням часових рядів або методами класифікації на основі статистичних даних [8]. За допомогою першої групи алгоритмів можна визначити прогнозування майбутніх значень в часовій послідовності, а другої - приналежність до певного класу. Зважаючи на специфіку дисертаційної роботи пов'язаної з двома станами (критичний чи не критичний) дефектів більш важливим буде використати методи класифікації машинного навчання.

Одним з розглянутих методів МН є дерева рішень, можна порівняти з процесом прийняття рішень людиною під час вирішення проблем. Дерева рішень використовуються для поділу на два чи більше класи об'єктів на основі набору правил і характеристик. Кожна гілка дерева представляє рішення, яке враховує певні умови або характеристики об'єктів. Дерева рішень особливо ефективні при роботі з категоріальними даними, тобто даними, які представлені у вигляді категорій чи класів. Такий метод може ефективно визначати правила і взаємозв'язки між різними категоріями. Як відомо для виявлення лінійних зв'язків між цільовими і незалежними змінними ефективніше використовувати лінійні моделі, такі як лінійна регресія чи метод опорних векторів (SVM) з лінійним ядром, аніж дерева рішень враховуючи їх нелінійність структури. Матриця невизначеності допомагає аналізувати помилки моделі і приналежності до класів такі як (Рис. 2.13):

- *True positive (TP)* - (істинно позитивні) кількість правильно визначених дефектів критичного стану

- *True negative (TN)* - (істинно негативні) кількість правильно визначених дефектів не критичного стану
- *False positive (FP)* - (хибно позитивні) кількість неправильно визначених дефектів критичного стану
- *False negative (FN)* - (хибно негативні) кількість неправильно визначених дефектів не критичного стану

Це важливо для того, щоб зрозуміти переваги та недоліки моделі та можливі шляхи її поліпшення.

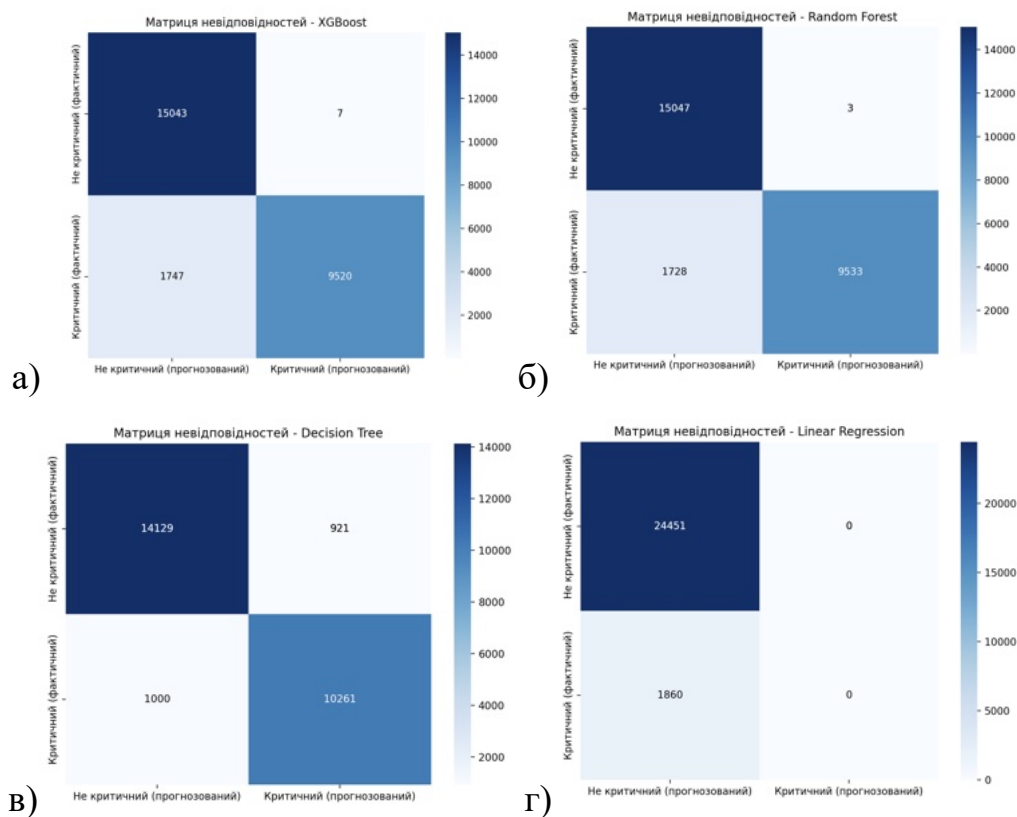


Рис. 2.13. Матриця невідповідностей фактичних і прогнозованих дефектів: а) - XGBoost , б) - випадковий ліс, в) - дерево рішень, г) - лінійна регресія

Також на основі отримати значення TP, TN, FP і FN з матриці невизначеності можна обчислити різні метрики ефективності моделі такі як чутливість, точність, специфічність, помилково негативний рейтинг та оцінка F1 [118]. Ця інформація відображає взаємодію між прогнозами моделі і дійсними мітками.

Точність ( $A$ ) є метрикою, яка визначає, яку частку вірно класифікованих елементів (або прогнозів) здійснила модель серед усіх елементів (прогнозів). Це відношення кількості правильно класифікованих елементів до загальної кількості елементів і обчислюється наступним чином:

$$A = \frac{TP+TN}{TP+TN+FP+FN}, \quad (2.1)$$

Рівень неправильної класифікації (error rate)

$$ER = \frac{FP+FN}{TP+TN+FP+FN}, \quad (2.2)$$

Ще однією метрикою є точність (precision ( $P$ )), яка обчислюється:

$$P = \frac{TP}{TP+FP}, \quad (2.3)$$

Показник повноти (Recall або True Positive Rate) показує кількість правильно прогнозованих дефектів

$$R = \frac{TP}{TP+FN}, \quad (2.4)$$

Специфічність (True Negative Rate (TNR)) вимірює відсоток правильно визначених негативних екземплярів відносно всіх дійсних негативних екземплярів і є результатом наступного співвідношення:

$$TNR = \frac{TN}{TN+FP}, \quad (2.5)$$

False Negative Rate (FNR) вимірює відсоток правильно визначених позитивних екземплярів відносно всіх дійсних позитивних екземплярів і обчислюється наступною формулою:

$$FNR = \frac{FP}{FP+TN}, \quad (2.6)$$

Ці метрики важливі для оцінки ефективності класифікаційних моделей містяться у Таблиці 2.5. Значення чутливості є важливий в ситуаціях, коли помилка визначення позитивного класу може мати серйозні наслідки (наприклад, у випадках техногенних катастроф). Точність дає загальну перевірку правильності моделі, а специфічність важлива в ситуаціях, коли помилка визначення негативного класу може бути критичною.

У цьому дослідженні використовується середня абсолютна помилка (*MAE*) і *F1* міра як критерії для оцінки ефективності моделі для різної глибини. Формули для параметрів *P* та *R* наведено в рівняннях вище (2.3) та (2.4) відповідно

$$F1 = \frac{2PR}{P+R}, \quad (2.7)$$

Таблиця 2.5. Результати оцінювання точності використання моделей машинного навчання для прогнозу змін станів дефектів на поверхнях

Модель	Гradientне підсилення	Випадковий ліс	Дерево рішень	Лінійна регресія
Точність (accuracy)	0.933	0.934	0.927	0.929
Рівень неправильної класифікації (error rate)	0.067	0.066	0.073	0.071
Точність (precision)	0.999	0.999	0.939	1
Повнота (recall)	0.896	0.897	0.934	0.929
TNR	0.999	0.999	0.918	0
FNR	0.104	0.103	0.066	0.071
F1	0.945	0.946	0.936	0.963

Основні метрики для оцінки точності регресійних моделей включають:

Середньоквадратична похибка (Mean Squared Error, MSE):

$$MSE = \frac{1}{n} \sum_{i=0}^n (y_i - \hat{y}_i)^2, \quad (2.8)$$

де *n* - кількість спостережень, *y<sub>i</sub>* - фактичне значення,  $\hat{y}_i$  - передбачене значення моделі.

Середня абсолютна похибка (Mean Absolute Error, MAE):

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (2.9)$$

де  $n$  - кількість спостережень,  $y_i$  - фактичне значення,  $\hat{y}_i$  - передбачене значення моделі.

Коефіцієнт детермінації (Coefficient of Determination,  $R^2$ ):

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.10)$$

де  $n$  - кількість спостережень,  $y_i$  - фактичне значення,  $\hat{y}_i$  - передбачене значення моделі,  $\bar{y}$  - середнє фактичне значення

Ці метрики визначають, наскільки добре модель пристосовується до фактичних даних.  $MSE$  та  $MAE$  використовуються для об'єктивної оцінки точності моделей регресії, що допомагає в рішенні прийняти щодо їхнього використання або подальшого вдосконалення. На рис. 2.14 показано що, як і  $MSE$ , чим менше  $MAE$ , тим більша точність.

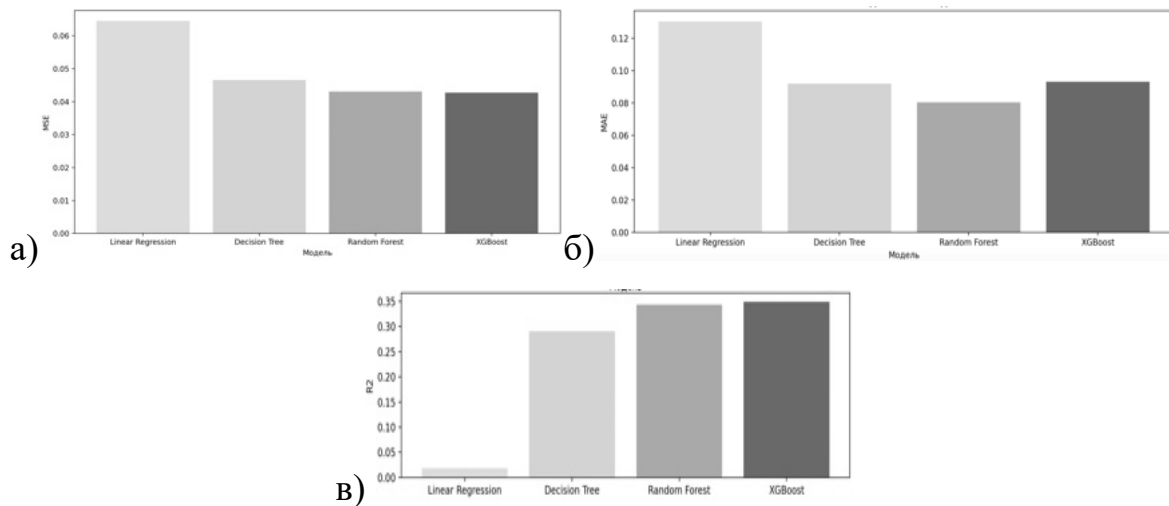


Рис. 2.14. Динаміка відхилень для різних алгоритмів а) - середньоквадратичного MSE, б) - середнього абсолютного MAE, в) коефіцієнта детермінації  $R^2$

Найменше меншу середню абсолютну помилку ( $MAE$ ) серед використаних моделей має градієнтне підсилення та випадковий ліс. Найвище  $R^2$  свідчить про те, що модель добре пояснює варіацію у вихідних даних.

## Висновки до розділу 2

У розділі побудовано та розглянуто статистичні моделі, а також історичні дані про аварії та ремонти елементів конструкцій інженерних об'єктів. Для аналізу та обробки зібраних даних про дефекти у цих об'єктах використано прогностичні моделі. Також, описано процес попередньої обробки даних розвідувального аналізу (тобто, виявлення пропусків та встановлення умов їх заповнення необхідною інформацією) з використанням методу вставлення найчастіших значень.

Розроблені прогностичні регресійні моделі, які є частиною створеного апаратно-програмного комплексу для оцінювання змін стану дефектів, можуть використовуватися для оптимізації регламентованого обслуговування дефектного елемента конструкції та розроблення системи прийняття рішень для оцінювання можливостей подальших змін.

Проаналізовано існуючі практичні рішення, що відображають основні підходи інженерії ознак для визначення важливості застосованих параметрів моделі та аналізу впливу факторів. Також, гіперпараметричне налаштування дає змогу покращити точність під час визначення важливості параметрів для цих алгоритмів. Визначення переваг і обмежень класифікаційних методів за результатами порівняльного аналізу допомагає зрозуміти, який метод може бути оптимальним для задачі прогнозу про зміни стану дефектів.

Обґрунтовано, що серед застосованих методів діагностування дефектів одним з найперспективніших є підхід градієнтного підсилення для обраного набору даних, зважаючи на отримані результати метрик точності.

Запропоновано формалізований підхід для відображення інформації про параметри зміни стану дефектів елементів конструкцій з урахуванням відібраних даних про попередні аварійні ситуації, що дає можливість оцінити показники ресурсу в певний момент часу.

Уведено модель оцінювання стану дефектів пошкодженого елемента конструкцій, яка спрощено відображає структуру параметрів причин та факторів ліквідації недоліків аварійного об'єкта і обґрунтовує зв'язки між ними.

Наведені вище уточнення стали підставою для поглиблення досліджень у напрямку діагностування дефектів, оскільки це допомагає у вдосконаленні стратегій технічного обслуговування та плануванні ресурсів для забезпечення тривалого та надійного функціонування конструкцій.

Проблемним аспектом роботи з даними є удосконалення технічних засобів та інформаційних технологій, що дозволяють впорядкувати обмін даними між компонентами різних підходів для розпізнавання дефектів та перевірки результатів прогнозування, які розглянуто у розділі 3.



### **РОЗДІЛ 3. ЗАСОБИ ВАЛІДАЦІЇ ВИКОРИСТАНИХ МЕТОДІВ ДЛЯ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ЗМІНИ СТАНУ ДЕФЕКТІВ У ЕІК**

Цей розділ присвячений опису апаратних та програмних методів та засобів для здійснення перевірки прогнозу про зміни стану дефектів на основі ідентифікованих пошкоджень ЕІК на зображенні. Для цього проведено покроковий опис компонентів, методів та засобів ІТ.

Основна ідея полягає в розробці системи розпізнавання дефектів для перевірки стану елементів конструкцій з метою покращення контролю та попередження можливих проблем у функціонуванні ЕІК. Для отриманих результатів розробки такої ІТ можна виділити наступні завдання:

1. дослідження, проектування та аналіз інформаційних потоків у формуванні сенсорних даних, методів надсилання їх з пристрою та обробці подій у веб сервісі зі здійсненням прогнозу в режимі реального часу.

2. виконання апаратного налаштування та обмін даними JetRacer AI Kit на базі мікрокомп'ютера Nvidia Jetson Nano для програмного комплексу розпізнавання дефектів елементів конструкцій;

3. розробка та імплементація клієнт-серверної взаємодії з метою перевірки прогнозу мікрокомп'ютером та моніторинговим веб середовищем;

4. доповнення існуючих підходів автоматизації прийняття рішень при дослідженні дефектів елементів конструкцій;

#### **3.1. Постановка задачі побудови системи розпізнавання для перевірки**

Для побудови системи розпізнавання дефектів елементів конструкцій можна сформулювати наступні основні етапи та функціональні вимоги:

- Вибір конкретних об'єктів (тріщини та корозійні процеси) для розпізнавання.
- Збір даних із здійсненням вимірювань та отримання зображень з поверхні об'єкта за допомогою камери чи іншого сенсора.

- Попередня обробка отриманих даних та аугментація даних для покращення роботи моделі.
- Створення моделі розпізнавання з вибором архітектури моделі та виконання тренування моделі на основі підготовлених даних.
- Валідація та тестування для перевірки ефективності моделі на відокремленому від тренувального наборі даних та тестування моделі на нових зображеннях з реальних об'єктів.
- Реалізація системи у вигляді веб-сервісу для взаємодії з іншими компонентами. Інтеграція із засобами відображення результатів (інтерфейс користувача).

### **3.2. Концепція прототипу інформаційно-діагностичної системи**

Проектування такої системи потребує виокремлення основних елементів ІТ, з яких варто врахувати наступні:

- Сенсори такі як камера для поверхневого та візуального розпізнавання або інші сенсори для збору даних з поверхні об'єктів та навколишніх факторів впливу на цей об'єкт.
- Модель навченої нейронної мережі або іншої моделі машинного навчання для виявлення дефектів на зображеннях. Під час розробки системи слід враховувати вимоги щодо точності розпізнавання, швидкодії, надійності та інші параметри, які визначають успішність реалізації.
- Групування та надсилання даних сучасними технологіями IoT, для цього методами комп'ютерного зору з використанням навчених моделей можна автоматизувати процес виявлення пошкоджень у різних матеріалах. Виділені в зображенні тріщини в кожному кадрі містять певну кількість пікселів, які використовуємо для оцінки стану об'єкту у певний час.
- Збережені дані у сховищі даних дозволяють проводити аналітику та діагностику на основі отриманих параметрів. Відповідні показники та

результати обробки відео з камери є важливими в контексті проведення перевірки прогнозування. А історичні дані про аварійність ЕІК дають змогу провести статистичний аналіз для визначення важливості вибраних інформативних параметрів та здійснити класифікацію даних.

- Інтерфейс користувача. Це засоби для взаємодії з користувачем, відображення результатів та управління системою.
- Алгоритми обробки даних. Алгоритми для аналізу та обробки даних перед їхнім використанням моделі.
- Система інтеграції. Механізми для взаємодії з іншими системами чи обладнанням.

Вимоги для інформаційної технології розпізнавання та прогнозування дефектів:

- Висока роздільна здатність зображень що дає змогу ефективно ідентифікувати дрібні дефекти;
- Ефективні алгоритми обробки зображень для виділення та аналізу дефектів;
- Забезпечення оптимального освітлення для отримання якісних зображень;
- Спроможність роботи системи в режимі реальному часі;
- Можливість виявлення різних типів дефектів (тріщини, корозія чи пітинги);
- Надійність та стабільність;
- Можливість інтеграції з іншими технологіями: Система повинна бути гнучкою та здатною інтегруватися з іншими технологіями для комплексного аналізу та прогнозування дефектів;
- Оцінювання стану дефектів у важкодоступних місцях конструкції може бути вирішене на основі автономних пристроїв (зокрема, мікрокомп'ютер Nvidia Jetson Nano, до яких приєднано додаткові сенсори вологи та інше).

Серед методів класифікації машинного навчання існують деякі відмінності в реалізації, які впливають на результат прогнозування. При введенні в запропоновану модель відомих входних параметрів (вік, кількість ремонтів, типу матеріалу) та зчитаних у певній ділянці конструкції ми отримаємо результати про можливі дефекти на поверхнях досліджуваного ЕК з урахуванням ймовірності зміни їх стану (критичний чи не критичний). Оскільки використовується мікрокомп'ютер, є можливість підключити сенсор DHT22, який додатково може надавати поточну інформацію про вологість і температуру навколишнього середовища. Крім цього, така система може розширюватися з додаванням сенсорів атмосферного тиску, вологості ґрунтів, термопари та інше. Технології Інтернету речей передбачають використання безпроводних технологій (Wi-Fi, Bluetooth, LTE) з використанням протоколу передачі даних до хмарного сховища даних.

На рис. 3.1 зображено загальну схему інформаційно-діагностичної системи під час відбору, обміну та аналізу даних щодо розпізнавання дефектів елементів конструкції.

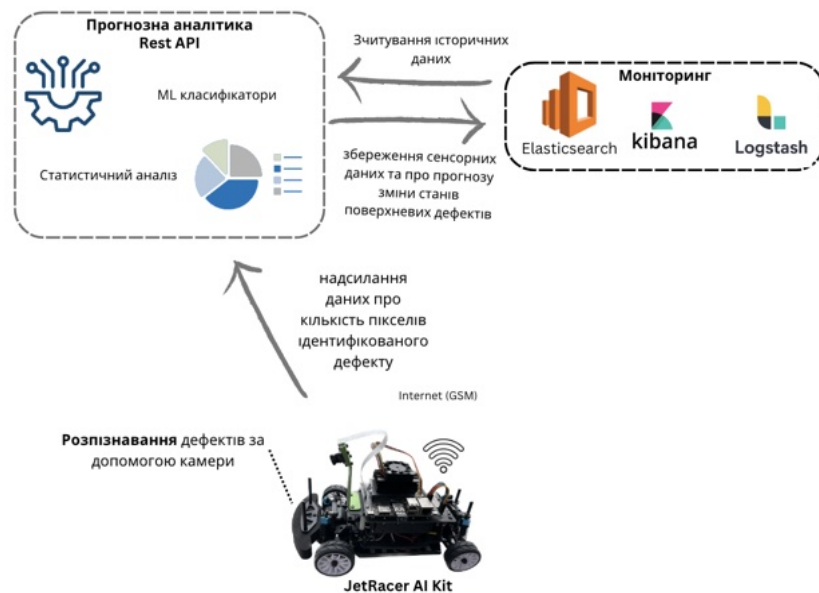


Рис. 3.1. Архітектура інформаційно-діагностичної системи для розпізнавання та прогнозування дефектів в елементах конструкцій.

Останнім кроком є збереження даних і моніторинг оброблених даних на інформаційній панелі. На противагу комерційним рішенням найпопулярніших хмарних середовищ Amazon Web Service (AWS), Microsoft Azure і Google Cloud Platform можна використати існуючі програмні інтерфейси Elasticsearch додавання великої кількості даних за одну операцію (Bulk API). Хоча хмарні сервіси мають багато переваг, таких як масштабованість, висока безпека, швидкість розгортання та доступність, вони також можуть мати свої обмеження такі як постійні витрати на тривалому використанні з великою кількістю даних [3]. Такі підходи в проектуванні ІТ зазвичай вигідні на початкових етапах, бо забезпечують користувацький інтерфейс для автоматизованого налаштування інфраструктури системи.

У роботі запропоновано зберігати дані в нереляційній базі даних Elasticsearch, враховуючи специфіку обміну даними на основі файлів JSON та правил IoT [5]. Такий підхід дозволить розгортати систему як і у хмарних середовищах так і для локальних застосувань. Інструментом для візуального перегляду обраної бази є інструмент Kibana та Logstash. Таким чином, інформаційна технологія може бути використана для прогнозування зміни стану з подальшою перевіркою під час розпізнаванні дефектів на зображеннях. Повний цикл обробки даних стосовно дефектів методами аналізу даних може бути використаний як концептуальне рішення для діагностики матеріалів.

### **3.3. Метод оптимізації передачі даних з IoT**

Поняття інформаційного потоку пов'язане з можливостями обміну даними між різними пристроями, сенсорами, системами чи хмарними сервісами.

У дисертаційній роботі розглянута одна з проблем IoT (оптимізація кількості інформації для передачі та отримання), а саме використання інформаційних потоків для обміну даними у ІТ. Зазвичай причинами недосконалості рішень є зменшення проміжних обчислень у пристрої, який

передає дані, через обмежений ресурс графічного (GPU) та центрального (CPU) процесорів. Тому розглянуто існуючі способи перетворення та передачі даних оброблених зображень та оптимізацію кількості даних для зменшення навантажень у ІТ.

В контексті Інтернету речей така взаємодія між фізичними об'єктами ґрунтується на обміні даних у мережі Інтернет через вбудовану логіку обробки та методів передачі даних у пристрої. Поняття інформаційного потоку використовується на різних етапах обробки даних: збору різноманітні дані (температура, вологість, положення чи зображення з камери), обробка та проміжні перетворення до певного формату, передача від пристрою до центрального сервера, зберігання історичних даних, аналіз для формування висновків чи прийняття рішень та надсилання відповіді для здійснення дій [105–107]. Така взаємодія може застосовуватися для різних задач починаючи від моніторингу параметрів навколишнього середовища до автоматизації технологічних процесів та управління діями пристроїв.

Для перетворення обробленого зображення у числове представлення з камери в ІТ до інших пристроїв або хмарних рішень можна використовувати різні методи і технології. Поряд з протоколами передачі даних прикладного рівня Open Systems Interconnection (OSI) моделі такими як: Hypertext Transfer Protocol (HTTP) або (Hypertext Transfer Protocol Secure) HTTPS, File Transfer Protocol (FTP), використовують популярні методи передачі даних у IoT: Message Queuing Telemetry Transport (MQTT), Constrained Application Protocol (CoAP), Advanced Message Queuing Protocol (AMQP). Для випадків, коли потрібне з'єднання з мобільними пристроями або IoT-сенсорами, можна використовувати Bluetooth, NFC або LAN для локальної передачі даних. Проте, вибір конкретного методу буде залежати від поставленого завдання, вимог до безпеки, швидкості передачі даних та особливостей роботи з пристроями та даними розглянутих у працях [108–111].

Результати досліджень, отримані в роботі [112–114], показали, що ефективність MQTT вища з точки зору простоти та безпеки, ніж у традиційних протоколів передачі даних. У виконаному дослідженні [5] показано можливість інтеграції запропонованої системи з протоколами передачі даних IoT.

Зазвичай під час розробки ІТ [115] надають перевагу формуванню запитів у вигляді байтів для пристроїв Field Program Gate Array (FPGA) або Raspberry Pi (Simulink за допомогою системного генератора Xilinx (XSG)). Зважаючи на можливість передачі даних меншу кількість даних, варто розглянути результати обробленого зображення. Крім цього, у праці [116] запропоновано здійснювати ідентифікацію об'єкта після отримання зображення із зчитуючого пристрою.

У працях [117] запропоновані методи нормалізації та фільтрування зображенням на основі колірних характеристик RGB під час застосування методів комп'ютерного зору для подальшої передачі цих даних. Для уникнення надмірної кількості надісланих повідомлень до сховища даних чи проміжного програмного інтерфейсу можна групувати ці дані на пристрої і передавати для подальшого аналізу зміни стану дефектів (наприклад, Bulk API в ES [118]).

Наприклад, зображення у чорно-білих тонах середньої роздільної здатності ( $512 \times 512$ ) потребує  $512 \times 512 \times 8$  біт для свого представлення. Тому, щоб зберігати та передавати цифрові зображення було практично, потрібно перетворювати в графічні формати стиснення зображень JPEG, PNG, GIF, WebP, HEIF/HEIC, BPG, AVIF. За допомогою такого методу надмірність зображень використовується для зменшення кількості бітів, необхідних для їх представлення. Натомість запропоноване рішення, надсилати цілочисельне значення кількості пікселів з IoT. У 16-розрядних операційних системах тип `int` зазвичай становить 16 біт або 2 байти. У 32-розрядних операційних системах тип `int` зазвичай становить 32 біти або 4 байти. Таким чином, спрощується система та зменшується навантаження на центральний сервер.

На підставі літературного огляду останніх розробок можна підсумувати, що процес конвеєру даних, зображений на рис. 3.2, є гнучким та може зазнавати модифікації, розширенню та адаптації як для розробки систем прийняття рішень в області дефектоскопії, так і для більш складних комерційних рішень.



Рис. 3.2. Схема конвеєру даних з IoT

Кількість бітів та об'єм інформації є ключовими показниками, які визначають ефективність передачі даних у комунікаційних системах. Чим більше кількість бітів або об'єм інформації, тим більше даних може бути передано протягом певного проміжку часу.

Кількість бітів ( $I$ ), які передаються через канал зв'язку обчислюється залежно від властивостей інформації (зокрема довжина та кількість повідомлень), яка передається програмному інтерфейсу і відповідає наступному співвідношенню:

$$I = L \cdot N, \quad (3.1)$$

де  $I$  - кількість біт,

$L$  - довжина кожного повідомлення (в бітах),

$N$  - кількість повідомлень.

Об'єм інформації ( $V$ ) може вказувати на загальну кількість інформації, яка передається протягом певного періоду часу.

$$V = I \cdot R, \quad (3.2)$$

де  $V$  - об'єм інформації (в бітах або байтах),

$I$  - кількість біт,



$R$  - частота передачі інформації (біти на одиницю часу, наприклад, біти за секунду).

Ефективність обміну інформацією в Інтернеті речей можна оцінити використовуючи вищезгадані показники для визначення споживаної пропускної здатності мережі та ресурсів. Також, важливо враховувати особливості мережі IoT, такі як затримка, об'єм інформації під час передачі даних, пропускна здатність (throughput), швидкість передачі даних (data transfer rate). Ці параметри ефективності системи показують інтенсивність інформаційних потоків та детальніше описані в розділі 4.1.

### **3.4. Доповнення методу розпізнавання дефектів результатами прогностичного аналізу**

Враховуючи різну “природу” даних для обох задач, що використовуються в цьому контексті, мають різні формати, типи або властивості, які варто підкреслити гібридність підходу. Це зумовлено тим, що модель із задачі розпізнавання дефектів доповнюють іншу з прогностичної аналітики стану дефекту. Реалізація в ІТ може бути виконана за допомогою інтеграції різних етапів обробки даних та використання відповідних моделей машинного навчання гібридного типу. Нижче наведено загальну блок-схему (рис. 3.3), яка включає етапи від збору даних до прийняття рішень. Вхідними даними є параметри стану навколишнього середовища з сенсорів та візуальні дані для розпізнавання дефектів. Оскільки у даних можуть бути викиди чи аномалії, то здійснюється перевірка на відповідність форматів та інших критеріїв. Якщо сенсорні дані можна надсилати одразу, то для перевірки прогнозу потрібне використання моделей машинного навчання для виявлення дефектів на основі отриманих даних. Для оцінки спостережень та моніторингу зміни станів дефектів отримані результати розпізнавання та інших важливих метаданих зберігаються у сховищі даних Elasticsearch для подальшого аналізу та візуалізації. Моделі можуть

містити реальні дані з сенсорів як тестовий набір даних і перевіряються з використанням класифікаторів для прогнозування змін у стані дефектів. На основі даних з розпізнавання можна здійснити перевірку прогнозу та провести аналіз змін у стані дефектів. Враховуючи результати розпізнавання та прогнозу аналітики можна здійснити прийняття рішення щодо обслуговування або ремонту обладнання. Додатково до цього, статистичне моделювання може застосовуватися для рекомендацій та стратегій, визначених на основі аналізу, для управління дефектами та підтримки обладнання. А зворотній зв'язок може використовуватися як збір додаткових даних та результатів взаємодії з системою для постійного вдосконалення моделей та стратегій.



Рис. 3.3. Алгоритм обміну інформаційними потоками у задачах розпізнавання та прогнозування

### 3.5. Апаратна архітектура ІТ для оцінювання ресурсу кіберфізичної системи “МІТ – СКЗ”

Перші кроки (рис. 3.2) після запуску пристрою реалізуються безпосередньо на мікрокомп'ютері (Raspberry PI/Nvidia Jetson Nano). Таким чином, забезпечується автономність роботи системи. За специфікою роботи із зображеннями та технологіями машинного навчання є більш ефективним Nvidia

Jetson Nano з GPU NVIDIA Maxwell з 128 ядрами CUDA, що дозволяє виконувати паралельні обчислення. Компактність цього мікрокомп'ютера дає змогу використовувати його у важкодоступних місцях з використанням невеликого споживання енергії. Модель цього пристрою підтримує інтерфейси: GPIO (General Purpose Input/Output), I2C, I2S, SPI, UART, що дозволяє його інтеграції з різноманітними сенсорами та пристроями IoT. У цьому мікрокомп'ютері є вбудована підтримка інфраструктури для машинного навчання, а саме програмні каркаси TensorRT і CUDA. Завдяки високій обчислювальній потужності та підтримці машинного навчання, Jetson Nano використовується для відеоаналітики та виявлення об'єктів на відеозаписах або в реальному часі. Можливість використання GPU для графічної обробки робить Jetson Nano ефективним у випадках, коли потрібно обробляти великі обсяги графічних даних, наприклад, для розпізнавання дефектів або аналізу відеопотоків. Інтеграція безпосередньо обчислень у Jetson Nano та IoT дозволяє здійснювати обробку даних та виконувати розподілені обчислення безпосередньо за допомогою мікрокомп'ютера, зменшуючи трафік в мережі та забезпечуючи більшу ефективність використання ресурсів.

Таблиця 3.1. Характеристика використаних мікрокомп'ютерів

<b>Характеристика</b>	<b>Raspberry Pi 4 (4 GB)</b>	<b>Jetson Nano (2 GB)</b>
Процесор	Broadcom BCM2711, Quad-core Cortex-A72	Quad-core ARM Cortex-A57 MPCore
Графічний процесор	VideoCore VI	NVIDIA Maxwell (128 CUDA cores)

Оперативна пам'ять (RAM)	4 GB LPDDR4	2 GB LPDDR4
USB порти	2 x USB 3.0, 2 x USB 2.0	4 x USB 3.0, 1 x USB 2.0
Відеовиходи	2 x micro HDMI (до 4K)	HDMI, DisplayPort
CSI (камери)	Так (2 x CSI)	Так (1 x CSI)
Модуль для машинного навчання	Ні	Так (NVIDIA CUDA, cuDNN)

Оскільки зображення складається з пікселів, роздільна здатність зчитувача відіграє важливу роль у зоні охоплення досліджуваної області. Крім того, часто об'єкти знаходяться в умовах обмеженого освітлення, наприклад, при визначенні місця дефектів труб або підземних комунікацій. Тому важливу роль тут відіграють можливості пристроїв на базі мікропроцесорів або мікрокомп'ютерів (зокрема, платформ JetRacer AI Kit і PiRacer AI Kit). Ці набори побудовані на основі мікрокомп'ютерів Raspberry Pi 4 та Nvidia Jetson Nano з приєднаними через інтерфейс Camera Serial Interface (CSI) камерами 5MP 1080P 160 ° 2560 × 1960 і HD Sony IMX219 8 MP 160 ° FOV 3280 × 2464 відповідно (Рис. 3.4). Застосування такого пристрою для виявлення дефектів перевірено у працях дисертанта [5, 13, 15]. Перевагою використання таких пристроїв є можливість обробки, збереження даних і зручність програмування безпосередньо в графічному середовищі пристроїв.

Системи розпізнавання на основі нейронних мереж демонструють достатню надійність і точність виявлення інформації шляхом поділу зображення

на частини та пошуку дефектів. Для цього були розроблені алгоритми процесу навчання нейронної мережі.



Рис. 3.4. Засоби для розпізнавання дефектів на основі мікрокомп'ютерів Nvidia Jetson Nano та Raspberry Pi

Для здійснення керування JetRacer Kit використано підходи аналогічні як у прикладі базової реалізації [119]. Введення параметрів дає змогу здійснювати управління на основі посилення рульового керування, зміщення рульового керування, канал рульового керування та інших.

### 3.6. Програмна архітектура для розпізнавання дефектів у кіберфізичній системі “МІТ – СКЗ”

На пристрої Nvidia Jetson Nano встановлена операційна система Ubuntu 18.04 LTS. Програмну архітектуру можна поділити на скриптову та головну програмну компоненту. Перша базується на написаних скриптових частинах, які починають роботу програмних реалізацій на етапі старту пристрою. Друга - реалізація безпосередньо зчитування вхідного потоку кадрів та виконання розпізнавання поверхневих дефектів елементів конструкцій (розділ 1).

Важлива структура файлів в архітектурі проекту, оскільки вона забезпечує чіткість та логічність організації коду (Рис. 3.5).

Всі файли моделей розміщені в підпапках *models/cnn/* та *models/yolo/*. Для моделі CNN в підкаталог *cnn* є файли *cnn\_model.pt*, *config.json* (Рис.3.6) та *onnx\_model\_name.onnx*. Для моделі YOLO в підкаталог *yolo* є файл *yolo\_model.pt*.

Дані для тестування, валідації та тренування розміщені відповідних підкаталог під *data/*. Файли *pytorch\_to\_onnx.py*, *utils.py*, *train.py* та *evaluate\_and\_send.py* розміщені безпосередньо в кореневій папці проекту.

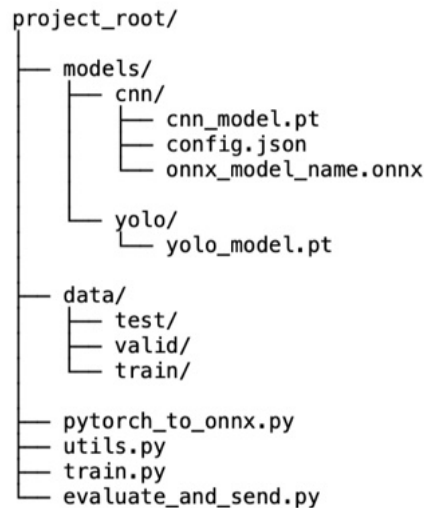


Рис. 3.5. Структура файлів для частини розпізнавання поверхневих дефектів

Спочатку потрібно ініціалізувати потік відеокамери для отримання відеосигналу. Потім здійснюється зчитування кадрів ітеративним способом із потоку відеокамери для подальшого аналізу. Наступним кроком відбувається завантаження попередньо навченої моделі для ідентифікації тріщин.

Після цього відбувається визначення загальної кількості кадрів у відеофайлі та зміна розміру вхідного зображення. Відкриття моделі та пропускання через неї кожен кадр. Використовуючи Argmax, визначаються мітку класу для кожного пікселя та відображаються відповідним кольором. Наступним кроком є зміна розміру маски та класу відповідно до розмірів вхідного зображення. Далі відбувається зважене поєднання вхідного зображення та маски для формування вихідної візуалізації. Після чого змінюється розмір вікна виводу та відображається результат ідентифікації тріщин.

```

{
  "settings": {
    "dataset_tags": {
      "train": "train",
      "val": "val"
    },
    "batch_size": {
      "train": 6,
      "val": 3
    },
    "data_workers": {
      "train": 3,
      "val": 0
    },
    "allow_corrupted_samples": {
      "train": 0,
      "val": 0
    },
    "special_classes": {
      "background": "bg",
      "neutral": "neutral"
    },
    "input_size": {
      "width": 256,
      "height": 256
    },
    "epochs": 3,
    "val_every": 0.5,
    "lr": 0.001,
    "momentum": 0.9,
    "lr_decreasing": {
      "patience": 1000,
      "lr_divisor": 5
    },
    "loss_weights": {
      "bce": 1.0,
      "dice": 1.0
    },
    "weights_init_type": "transfer_learning",
    "validate_with_model_eval": true,
    "gpu_devices": [
      0
    ]
  },
  "out_classes": [
    {
      "title": "crack_bitmap",
      "shape": "bitmap",
      "color": "#00021B"
    },
    {
      "title": "bg",
      "shape": "bitmap",
      "color": "#222222"
    }
  ],
  "class_title_to_idx": {
    "bg": 0,
    "crack_bitmap": 1
  }
}

```

Рис. 3.6. Файл налаштування навчальних параметрів у форматі JSON

Для налаштування процесу навчання розроблений конфігураційний файл (config.json на рис. 3.5) з функцією `get_default_config()`, яка повертає словник з параметрами конфігурації за замовчуванням для навчання моделі на деякому наборі даних:

- *dataset\_tags*: визначає теги для різних частин набору даних, таких як 'train' для тренувальної та 'перевірка' для валідаційної вибірок.
- *batch\_size*: задає розмір пакету даних (batch size) для тренування та валідації. Наприклад, 6 для тренування та 3 для валідації.
- *data\_workers*: кількість робочих процесів, які використовуються для завантаження даних для тренування та валідації 3.
- *allow\_corrupted\_samples*: показує, чи дозволені пошкоджені зразки у тренувальній та валідаційній вибірках.
- *special\_classes*: визначає особливі класи, такі як 'background' та 'neutral', і присвоює їм відповідні теги.
- *input\_size*: задає розмір вхідного зображення для моделі
- *epochs*: Кількість епох для тренування моделі
- *val\_every*: частота проведення валідації під час тренування (наприклад, кожні 0.5 епохи).

- *lr*: початковий крок навчання (learning rate).
- *momentum*: параметр моменту для оптимізатора.
- *lr\_decreasing*: параметри для зменшення швидкості навчання (learning rate).
- *loss\_weights*: ваги для різних компонентів втрат (бінарна крос-ентропія, dice).
- *weights\_init\_type*: тип ініціалізації ваг моделі (TRANSFER\_LEARNING або CONTINUE\_TRAINING).
- *validate\_with\_model\_eval*: перевіряє чи використовувати оцінку моделі для валідації.
- *gpu\_devices*: вказує, які GPU пристрої використовувати для навчання (у цьому випадку, лише один пристрій, GPU з індексом 0).

### 3.6.1. Структура сервісної частини прогнозування для дослідження кіберфізичної системи “МПТ – СКЗ”

Основною мовою програмування для реалізація серверної частини обрано Python версії 3.10. Вибір мови програмування зумовлений тим, що досить легко інтегрується з бібліотеками для розгортання на програмному каркасі FastAPI та роботою з моделями машинного навчання та використанню інтелектуальних систем. Для розгортання веб сервісної частини клієнт-серверної архітектури взаємодія здійснюється обробкою клієнтських запитів, які можуть включати визначені маршрутні функції з відповідними обробниками подій та сервера на основі FastAPI з підтримкою швидкого асинхронного веб-сервера Uvicorn. Оскільки задача прогнозування повинна виконувати тренування моделі, прогнозування та розгортання, то аналогічні обробники подій визначені для виконання конкретних дій та управління функціональністю системи (розділ 2).

Інтегрування між компонентами через API часто використовуються для моделей машинного навчання або інших прогнозуючих системах. Таку взаємодію між частинами каркасу для візуального представлення зв'язків та



обміном інформацією показано на діаграмі послідовностей (рис. 3.7). Swagger UI дозволяє легко документувати та взаємодіяти з API, що полегшує його розробку та використання [120, 121], автоматизація з Docker [122] та розширеними функціями [123].

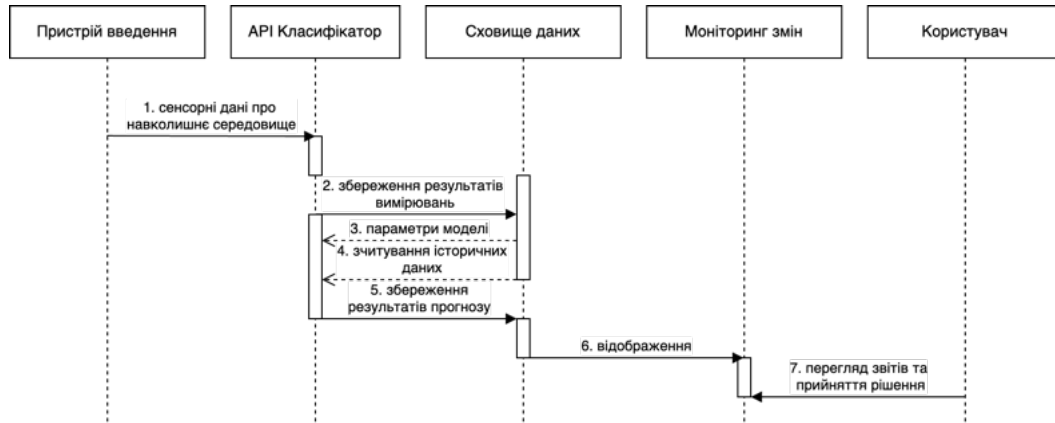


Рис. 3.7. Діаграма послідовностей роботи класифікатора

На етапі проєктування клієнт-серверного каркасу при виборі інструментів важливо враховувати протоколи та можливість взаємодії, відкритість та доступність технологій, оптимізацію часу розробки та відслідковування змін в режимі реального часу, забезпечуючи ефективну та надійну інфраструктуру для розробки IT. Загалом структуру такого каркасу можна поділити на наступні рівні:

- *рівень даних* в клієнт-серверній архітектурі може включати в себе використання Object-Document Mapping (ODM) для зручної взаємодії з нереляційними базами, обробку та обмін даними у форматі CSV для простоти інтеграції, а також використання моделей даних в Python для роботи з різноманітними джерелами та представлення інформації.
- *У рівні представлення інформації* на користувацькому інтерфейсі в клієнт-серверній архітектурі використання Kibana, що може забезпечити потужні можливості для створення динамічних та візуально збагачених графічних елементів та діаграм. Для більш гнучкої візуалізації можна використати D3.js (аналогічно як у роботі [2]) спільно з Django, як серверним фреймворком, це може забезпечити повністю функціональний та

ефективний інструментарій для створення веб-інтерфейсу. Додатково, інтеграція графіків моніторингу параметрів та генерація звітів може забезпечити засоби для візуалізації та аналізу динаміки даних, а також для зручної звітності з параметрів системи [124].

- У рівні бізнес-логіки класифікації дефектів з використанням бібліотеки `scikit-learn` можна реалізувати потужні алгоритми машинного навчання для автоматизованої ідентифікації та класифікації дефектів. Крім цього, інтеграція з веб-сервісом та використання маршрутних функцій може забезпечити доступ користувача до функціональності системи через API. Додатково до класифікації дефектів з метою здійснення прогнозу, можна розглядати можливості додавання аналітики, обробки подій та логіки для автоматичної обробки та реагування на виявлені дефекти.
- Рівень доступу реалізовує взаємодії з іншими компонентами каркасу на основі бібліотек до бази даних `Elasticsearch` з формуванням запитів `Domain-Specific Language (DSL)` [125], `Bulk API` або `Beats` [126], візуалізаційний інтерфейс `Kibana` та інструмент обробки журнальних подій `Logstash` [127].

Описані рівні можна зобразити в діаграмі компонент (рис. 3.8), що показує взаємодію між кожною частиною. Крім програм і засобів комунікації в діаграмі компонент показано організацію з'єднання між сховищем даних, веб сервісом, для прогнозування, інтеграційні точки з пристроями та засоби візуалізації отриманої інформації. Ці частини апаратно-програмного комплексу можуть взаємодіяти одна з одною через різні механізми комунікації, такі як мережа, API, міжпроцесорна взаємодія тощо. Основним компонентом комунікації є веб-інтерфейс за допомогою якого користувач може взаємодіяти з усією системою.

структурні та логічні зв'язки між усіма компонентами програмного забезпечення варто організувати за допомогою існуючих підходів до проектування відповідних систем, тобто архітектурних патернів.

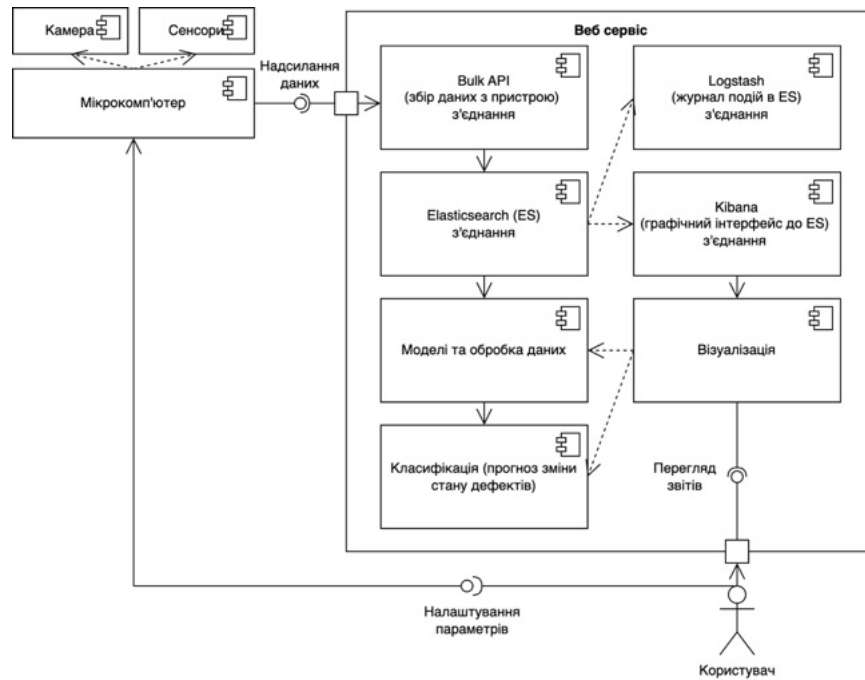


Рис. 3.8. Діаграма компонент сервісної частини

Для описаних рівнів каркасу можна застосувати варіацію архітектурного патерну *Модель-Вид-Контролер (MVC)* - *Модель-Вид-Шаблон (MVT)*, який є адаптований для веб-додатків та застосовується для опису його компонентів: моделей для структури даних, шаблону для візуалізації, і виду для бізнес логіки. Для реалізації *Модель-Вид-Шаблон (MVT)* для обробки даних з розпізнавання дефектів як клієнта так і сервера для прогнозування на основі історичних даних можна застосувати до рівнів каркасу.

Основні етапи реалізації:

- *Модель (Model)*: Створена модель для зберігання історичних даних про дефекти та інші відомості, які потрібні для прогнозування.
- *Вид (View)*: Створений вид, який буде відповідати за обробку запитів, отримання історичних даних, виклик моделі для побудови прогнозу та відображення результатів.
- *Шаблон (Template)*: Створений HTML-шаблон для відображення результатів прогнозу.

- *Бізнес-логіка та обробка запитів*: Реалізована функція або модуль для прогнозування дефектів на основі історичних даних.

Також додаткова логіка для збору нових даних, обробки введення від користувачів і т. д.

У описаній схемі (рис. 3.9) використання HTTP/HTTPS-запитів дозволяє взаємодіяти клієнтам (реалізовані на пристроях, таких як JetRacer AI Kit або Raspberry Pi) з серверною частиною системи.

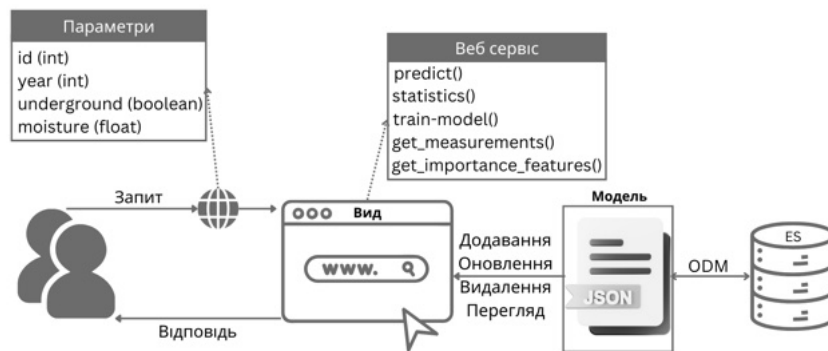


Рис. 3.9. MVC архітектура розробленої сервісної частини

“Вид” є компонентом взаємодії з користувачем на основі запитів та відповідей і може бути реалізована через веб інтерфейс або маршрутні функції веб сервісу. Основні дії відбуваються з моделлю на основі HTTP/HTTPS методів, які відповідають наступним операціям: додавання, перегляд, видалення, оновлення даних. Дані зберігаються у базі даних, а взаємодія її з моделлю реалізується через Object-Document Mapping (ODM). Враховуючи, що для обробки даних в режимі реального часу та інтерактивних прогнозуючих систем важливим є низька латентність запитів, різноманітність типів даних (текст, JSON, графові дані, геопросторові дані та інші) та швидкість для запису та оновлення даних з великими інформаційними потоками. Більшість з них мають сферу застосування у задачах, бо серед популярних рішень є MongoDB (для документ-орієнтованих даних), Cassandra (для розподіленого зберігання даних), Elasticsearch (для пошуку та аналізу даних) та інші.

Оскільки Elasticsearch є розподіленою системою збереження документів, і відноситься до поняття "документів" замість "реляційних таблиць". Сховище даних може реалізовуватись у стеку суміжних технологій EKL (Elasticsearch, Kibana, Logstash).

У веб сервісі реалізовані відповідні функції з'єднаних з базою даних для виконання операцій перегляду та додавання з можливим розширенням функціоналу, а саме видалення та оновлення даних (Рис. 3.10).



Рис. 3.10. Маршрутні функції

Розроблений веб сервіс містить наступні маршрутні функції (рис. 3.10):

- *statistics* призначений для обробки HTTP/HTTPS-запитів GET для отримання статистичної інформації. Може повертати статистичні дані або сумарну інформацію, пов'язану з застосунком.
- *classify* призначений для обробки HTTP/HTTPS-запитів POST для проведення класифікації. Може включати надсилання даних для класифікації, і сервер відповідає результатом класифікації.
- *get\_feature\_importance* призначений для отримання інформації про важливість ознак. Може повертати відомості про значущість чи важливість різних ознак у даних.

- *train-model* призначений для обробки HTTP/HTTPS-запитів POST для тренування моделі машинного навчання. Відповідає за надсилання даних для тренування на сервер, який потім тренує модель на основі цих даних.
- *get-measurement* з вказаним *item\_id* призначений для обробки HTTP/HTTPS-запитів GET для отримання конкретних вимірювань, ідентифікованих за *item\_id*. Змінна *item\_id* використовується для отримання відповідного вимірювання.
- *get-measurements* призначений для обробки HTTP/HTTPS -запитів GET для отримання списку вимірювань.
- *generate-report* призначений для генерації звіту.

### 3.6.2. Програмні компоненти для дослідження кіберфізичної системи “МІТ – СКЗ”

Програмні компоненти у контексті веб серверної архітектури складаються з трьох частин, що дозволяє краще організувати та взаємодіяти з різними аспектами системи. Синхронізація дій та обмін даними між складовими частинами у ІТ виконуються через HTTP/HTTPS запити (Рис.3.11).

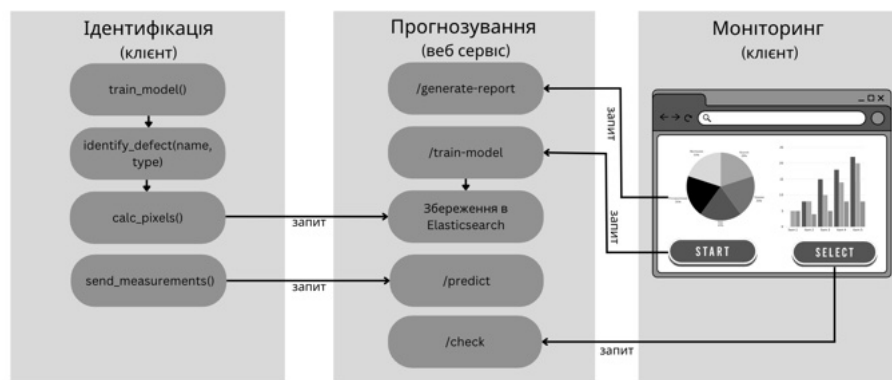


Рис. 3.11. Діаграма активностей для потоків та обміну даними

Основний потік інформації від клієнтів до веб сервісу реалізується наступним чином:

- *train\_model()* метод викликається для розпізнавання та передає дані методу *identity\_defect(name, type)*. Відповідні дані використовуються для

ідентифікації дефектів в об'єктах або для навчання моделі для подальшого використання.

- *send\_measurements()* метод після розпізнавання надсилає вимірювання або вихідні дані до */predict* маршрутної функції API, які необхідні для подальшого аналізу.
- */predict* маршрут API призначений для обробки отриманих вимірювань. Клієнт викликає та взаємодіє з певною функцією чи сервісом, яка використовує навчену модель для класифікації дефектів або інших характеристик.
- */generate-reports* маршрут API використовується для створення звітів на основі результатів збережених на етапі прогнозування */predict*. На цьому етапі завантажуються додаткові дані з сховища даних, щоб доповнити інформацію, а потім обробляються ці дані відповідним чином, створюючи представлення інформації у форматі JSON для візуалізації результатів.

Веб сервіс отримує та надсилає інформацію через вищезгадані маршрутні функції до бази даних. Зважаючи на відкритий код та суміжні інструменти (Bulk API, Kibana, Logstash), підтримку структурованих та неструктурованих даних та можливості повнотекстового пошуку інформації обрано нереляційну базу даних Elasticsearch 8.12.1 [128]. Оскільки база даних реалізована на основі технології Apache Lucene, то можна сформулювати запити різної структури такі як: географічні, складені, рядкові, діапазонні та інше. Дані зберігаються у відповідних індексах, а операції маршрутизації та ребалансування виконуються автоматично при додаванні нових документів (Рис. 3.12 та Рис. 3.13). Надійність описаного підбору засобів збереження інформації ефективніше при відновленні даних [8] та забезпечують високу доступність з додатковими дублюючими джерелами інформації [11].

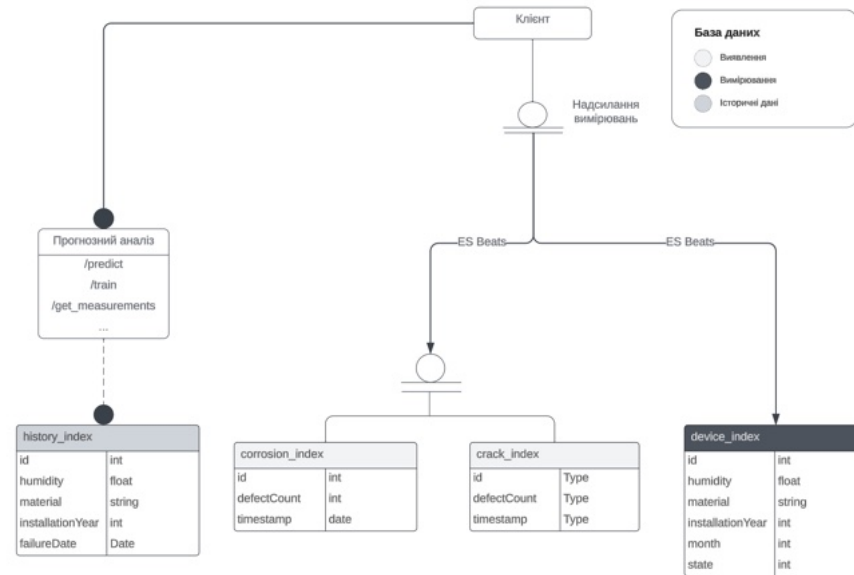


Рис. 3.12. Діаграма сутностей та зв'язків між програмними компонентами



Рис. 3.13. Список збережених даних з інформацією у індексах Elasticsearch: а) зчитані сенсорні дані, б) історичні дані, в) про корозійні стани дефектів на поверхні, г) про стани дефектів типу тріщин на поверхні



У Kibana є можливість візуалізувати різні параметри наприклад на Рис. 3.14 та Рис. 3.15 показано найчастіше використані 10 значень *defectCount* в індексі *corrosion\_index*.

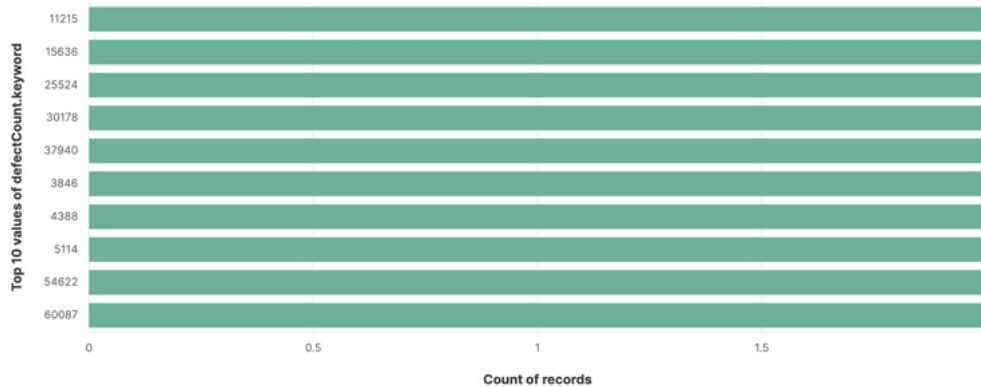


Рис. 3.14. Найчастіше використані 10 значень *defectCount* в індексі *corrosion\_index*

Type	Name	Documents (%)	Distinct values	Distributions	Actions
t	defectCount	1,000 (100%)	994		
k	defectCount.keyword	2,003 (100%)	1982	 top 10 of 1982 categories	

DOCUMENTS STATS		TOP VALUES	
count	2003	11215	2 (< 0.1%)
percentage	100%	15636	2 (< 0.1%)
distinct values	1982	25524	2 (< 0.1%)
		30178	2 (< 0.1%)
		37940	2 (< 0.1%)
		3846	2 (< 0.1%)
		4388	2 (< 0.1%)
		5114	2 (< 0.1%)
		54622	2 (< 0.1%)
		60087	2 (< 0.1%)
		Other	1983 (99%)

Calculated from 2,003 records.

Type	Name	Documents (%)	Distinct values	Distributions	Actions
#	timestamp	2,003 (100%)	12	 1,708,07... 1,708,08...	

Рис. 3.15. Аналіз значень атрибутів у Kibana

У моніторингових системах візуалізація інформації дозволяє швидко реагувати на зміни та проводити аналіз [2]. Надійність програмного інтерфейсу у клієнт-сервісній архітектурі доцільно перевіряти за допомогою аналізу відповідей на запити з позитивними та негативними сценаріями [10]. Крім цього,

при цьому варто визначити граничні навантаження (розділі 4.1) на програмний інтерфейс і забезпечити перемикання між джерелами даних [11].

### 3.7. Процес та результати валідації моделей

Демонстрація роботи схеми потоків та обміну даними (рис. 3.11) починається з етапу навчання (розділ 3.5). Після цього на основі попередньо натренованих моделей може здійснюватися ідентифікація дефектів за типом (зокрема, корозія чи тріщина). На цьому етапі відбувається паралельна обробка запитів через надсилання даних про ідентифіковано кількість пікселів та показники вологості, які можуть впливати на прогноз зміни стану дефектів на поверхнях. Такі запити виконуються через доступні можливості надсилання наборів даних за допомогою Bulk API (рис 3.16).

```
cURL ▾
1 curl --location 'http://localhost:9200/corrosion_index/_bulk' \
2 --header 'Content-Type: application/json' \
3 --data '{ "index" : { "_index" : "corrosion_index", "_id" : "1" } }' \
4 { "defectCount": "61356", "timestamp": 1708070493}
5 { "index" : { "_index" : "corrosion_index", "_id" : "2" } }
6 { "defectCount": "60835", "timestamp": 1708070495}
7 { "index" : { "_index" : "corrosion_index", "_id" : "3" } }
8 { "defectCount": "59253", "timestamp": 1708070497}
9 { "index" : { "_index" : "corrosion_index", "_id" : "4" } }
10 { "defectCount": "30453", "timestamp": 1708070499}
11 '
```

Рис. 3.16. Приклад curl команди з Bulk API для надсилання даних про кількість пікселів

Після відповідної обробки даних вони зберігаються у сховищі даних (Рис.3.13в). Паралельно до цього процесу відбувається зчитування даних з сенсора вологи DHT22 та формування тіла запиту:

```
{ "humidity": 20, "installation_year": 1990, "material": "steel",
  "current_date": "1685096635" }
```

Значення вологості додається до попередньо заданих відомих даних в маршрутну функцію `/predict` з відповідними параметрами (рис. 3.17)

predictions

POST /predict Run Prediction

Parameters

Name	Description
material * required string (query)	steel
installation_year * required integer (query)	1990
current_date * required string (query)	1685096635
humidity * required integer (query)	20

Execute Clear

Responses

Рис.3.17. Маршрутна функція */predict* з параметрами

Результатом є відповідь з програмного інтерфейсу, а оцінка *confidenceNegative* і *confidencePositive* базується на матриці невідповідностей на основі якої і *predictedLabel* із значеннями “*Defect*” (критичний) чи “*No defect*” (не критичний) (Рис. 3.18).

Responses

Curl

```
curl -X 'POST' \
  'http://127.0.0.1:8000/predict?material=steel&installation_year=1990&current_date=1685096635&humidity=20' \
  -H 'accept: application/json' \
  -d ''
```

Request URL

```
http://127.0.0.1:8000/predict?material=steel&installation_year=1990&current_date=1685096635&humidity=20
```

Server response

Code	Details
200	<p>Response body</p> <pre>{   "prediction": [     {       "material": "steel",       "year": 1990,       "date": "1685096635",       "humidity": 20,       "predicted_label": "No defect",       "confidenceNegative": 0.9939324249291,       "confidencePositive": 0.604685758789     }   ] }</pre> <p>Response headers</p> <pre>content-length: 185 content-type: application/json date: Tue, 23 Jun 2024 08:13:36 GMT server: uvicorn</pre>

Рис.3.18. Відповідь маршрутної функції */predict*

Таким чином, на основі збережених даних про timestamp (позначка часу) можна оцінити зміни станів дефектів на поверхнях, використовуючи візуалізацію змін значень параметрів вологості з прогнозованим станом дефекту (*state*) та кількості пікселів ідентифікованого дефекту (*defectCount*). Особливо такий підхід

може бути корисний для здійснення діагностики з періодичними перевірками стану ЕІК.

У системах можуть бути як переваги, так і недоліки, які впливають на їхню ефективність та функціональність. Переваги і недоліки у системах зазвичай зумовлені різними чинниками, такими як технології, використані матеріали та специфікації системи. Переваги запропонованих розробок можуть бути пов'язані з підвищенням продуктивності, ефективністю використання ресурсів, покращення надійності чи зручності в їх експлуатації. З іншого боку, недоліки можуть бути пов'язані з масштабованістю, складністю обслуговування чи залежність від зовнішніх чинників.

### 3.8. Переваги, обмеження та подальший розвиток ІТ

Перевагою такого підходу є концептуальне рішення формування та передачі інформаційних потоків у задачах розпізнавання та прогнозування.

Серед технічних недоліків JetRacer Kit має обмеження у розмірах (рис. 3.19), що унеможливує використання його у менших, аніж трубу менше ніж 15 сантиметрів у діаметрі, тим більше з її вигинами під гострими кутами.

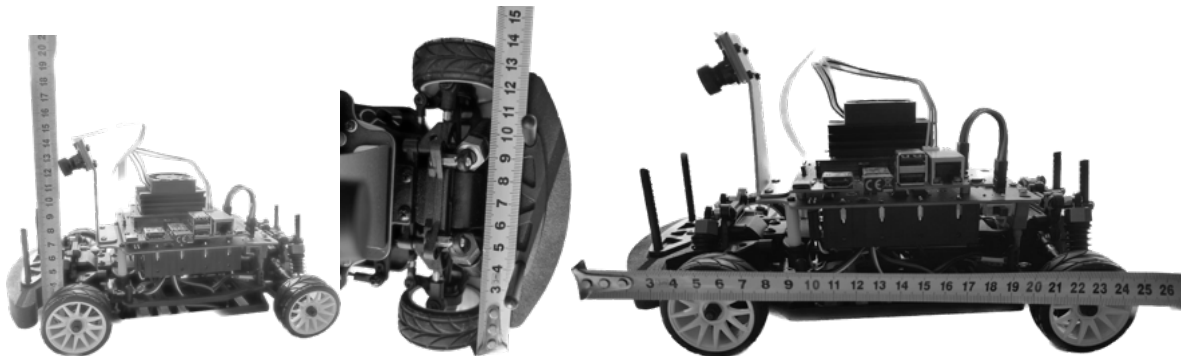


Рис.3.19. Розміри JetRacer Kit

У відкритому просторі чи невеликих ділянках труби таке рішення може використовуватися цілком успішно. Проте обмежений інтернет сигнал в трубі (замкнутий простір) може стати завадою для надсилання даних. Тому

пропонується здійснювати усю трансформацію і запис у файли. Тільки при наявності Інтернет з'єднання будуть поступово передаватися дані з затримкою. Для уникнення перевантаження системи дані можуть надсилатися впродовж певного часу з створеної черги запитів (queue message). Деякі аспекти вирішення таких проблем розглянуті у розділі 4. Використання дротового з'єднання ускладнює проходження пристрою та сповільнює через надмірну кількість пристроїв-посередників між засобом збору інформації та центральним сховищем даних, проте може бути надійнішим підходом для збору інформації [44].

Одним із напрямків подальших досліджень можна розглянути розпізнавання різних типів дефектів на основі визначених колірних характеристик. Окрім того, планується дослідження змін станів дефектів у часі на основі значень кількості пікселів ідентифікованих дефектів на зображеннях до часових проміжків зібраних протягом кількох діагностик ЕІК. Запропоновані методи та методики щодо методів та засобів інтелектуального аналізу даних є важливими для виконання відбору та більш детального опрацювання даних.

Однією з розглянутих альтернативних комерційних сервісів є Amazon Web Service Rekognition розглянуті у роботі [3], який можна використовувати для розпізнавання на основі завантаженого зображення з маркованими наборами даних перед навчанням. Недоліком такого підходу є обробка великих обсягів даних, що може призвести до фінансових витрат. Тому у процесі прийняття рішень варто звернути увагу на бюджетні обмеження, умови використання кожного конкретного сервісу та інші фактори.

У роботі можна здійснити розширення з використанням GPS модулів для передачі поточного розташування пристрою, що дає змогу здійснювати географічний аналіз дефектів в області дослідження.

### Висновки до розділу 3

На даний момент аспекти узгодженості між результатами розпізнавання дефектів та інформацією щодо подальшого прогнозу зміни стану дефектів не є остаточно розв'язаною і потребує окремих досліджень для конкретних даних.

Процеси розпізнавання та прогнозування дефектів у елементах конструкцій потребують сучасних рішень і для цього використовуємо інформаційні технології (ІТ).

Розроблено ІТ, яка в контексті гібридного підходу дає змогу поєднати розпізнавання дефектів ЕІК та прогнозування зміни їх стану на основі комплексного підходу відповідно до введених вхідних параметрів. Зв'язок аспектів розпізнавання та прогнозування і реалізований обмін даними між двома частинами програмно-апаратного комплексу дозволяє перевірити прогноз навченої моделі з урахуванням використаних реальних даних. Критерієм ефективності прогнозування є точність класифікаторів вибраного і найкраще підготовленого для набору даних дерева рішень. Опрацьовано вибірку даних із відкритих джерел стосовно зміни стану дефектів і для контролю відповідних процесів використано критерії якості. В результаті маємо поєднання засобів і методів і це дає переваги реалізації швидкого зв'язку та забезпечує умови автономної роботи відповідної кіберфізичної системи. Крім того, відзначено важливість процедури впливу поточних інформативних параметрів на прогнозування поведінки дефектів та розробку методу оптимізації передачі результатів аналізу зображень за допомогою підходів ІоТ. Відповідна технологія може бути розширена за рахунок використання додаткових контролюючих сенсорів і методів інтелектуального аналізу даних.

Окрім цього, розглянуті підходи до архітектури глибокого навчання UNet, а саме, до формування оптимальної кількості шарів енкодера та декодера для покращення ефективності методів оцінювання ресурсу кіберфізичної системи “металевий підземний трубопровід (МПТ) – СКЗ”.

Бездротові системи в контексті Інтернету речей дають змогу впроваджувати технічні рішення, для яких характерна певна автономність. Під час бездротової передачі даних у мережі Інтернет важливою характеристикою є час обміну інформацією. Оптимізацію обсягу даних запропоновано здійснювати завдяки заміні байтової послідовності значень кожного кадру на числові значення із зазначенням кількості пікселів в обробленому зображенні дефекту. Для зменшення навантаження на центральний сервер розроблено відповідну ІТ і завдяки цьому усі проміжкові обчислення доцільно проводити безпосередньо на мікрокомп'ютері. Наведено приклади, які підтверджують необхідність та доцільність розробки методів оптимізації інформаційних потоків під час передачі даних, які спрямовані до центрального сервера для зберігання і аналізу станів дефектів.

Проведення інтелектуального аналізу даних стосовно зображень дефектів на поверхнях елементів конструкцій дозволяє удосконалити метод оцінювання ресурсу кіберфізичної системи. Відповідні результати викладено у наступних працях дисертанта [2, 3, 8, 13, 14, 15].

Виконані дослідження надали можливість виокремити компоненти кіберфізичної системи, описати взаємодію між ними та навколишнім середовищем, сконцентрувати увагу на структуру та функції кожної частини. Опис компонентів кіберфізичної системи є важливим етапом для подальшого розроблення та оптимізації методик розпізнавання дефектів ЕІК, а також для виявлення можливостей для поліпшення їх функціональності та продуктивності.

## **РОЗДІЛ 4. МЕТОДИ ОЦІНЮВАННЯ ДІАГНОСТИКИ ТА АНАЛІЗУ ІНФОРМАЦІЙНИХ ПОТОКІВ**

Створено математичну та функціональну модель, яка акумулює великий обсяг інформації, і запропоновано відповідну концептуальну схему розробки інформаційних технологій для аналізу інформаційних потоків в руйнівних процесах у елементах конструкцій. У цьому розділі показано використання підходів інформаційних технологій та інженерії якості для підвищення ефективності підземних металевих трубопроводів (ПМТ) на основі дослідження множини інформативних параметрів контролю як технічного стану так і режиму роботи газопровідних систем з урахуванням катодного захисту (КЗ) та ризиків. Враховуючи важливість синхронізації між процесами обчислення кількості пікселів ідентифікованого дефекту та прогнозу зміни станів поверхневого дефекту, варто розглянути навантаження на ці частини для мінімізації затримок в обробці даних, що може вплинути на коректність інтелектуального аналізу даних.

Тому проаналізовано основні критерії моніторингу критичних показників навантаження при збереженні даних, що дозволяють оцінити інтенсивність інформаційних потоків.

Метою даного розділу є використання підходів до розробки інформаційних технологій та інженерії якості для підвищення ефективності підземних металевих трубопроводів (ПМТ) на основі дослідження множини інформативних параметрів контролю як технічного стану так і режиму роботи газопровідних систем з урахуванням катодного захисту (КЗ) та ризиків.

### **4.1. Постановка задачі оцінки характеристик інформаційних потоків**

Ефективність використання технологій глибокого навчання для розпізнавання дефектів розглянуто у розділі 1. Описані особливості точності та швидкості обчислення кількості пікселів дають змогу оцінити переваги та



недоліки різних алгоритмів. У розділі 2 проаналізовано точність прогнозування та основні фактори, що впливають на зміну стану дефектів. Оскільки прогнозування станів дефектів реалізоване у REST API, то такі параметри ефективності як пропускну здатність та частоту помилок перевірені у праці [11]. Запропоновані методи і засоби для інтеграції обох частин інтелектуального аналізу дефектів описані у розділі 3.

Синхронізація потоків даних з двох частин (прогнозування змін станів дефекту та валідації відповідно до кількості пікселів ідентифікованого дефекту) важлива для коректності оцінювання змін станів дефектів на поверхнях.

В контексті систем прийняття рішень та обміну інформацією взаємодію між інформаційними потоками в процесі моделювання можна представити як систему масового обслуговування [129].

Під час інтенсивного обміну інформацією в процесі обробки великої кількості даних важливу роль відіграє інтенсивність вхідного потоку  $I$ . При великій кількості запитів запити обробляються у порядку черги за принципом “перший прийшов - перший вийшов” [130]. У запропонованому IoT підході затримка може бути пов'язана з відсутністю зв'язку. Проте ця проблема вирішується накопиченням відповідних інформації та надсилання даних групами. Ще одним аспектом формування інформаційних потоків з затримкою передачі інформації можна вважати недостатню потужність на одному з вузлів системи при обміні даних.

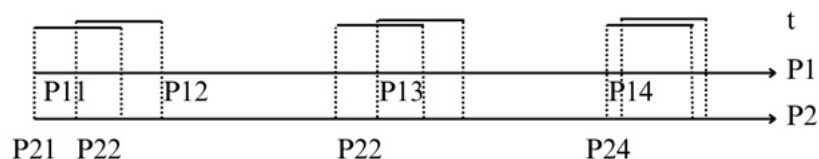


Рис. 4.1. Приклад формування потоків

де  $P1$  та  $P2$  - це ймовірність подій на певному проміжку,  $t_i$  - моменти входу інформацій в сектор обслуговування,  $t_0$  - момент виходу,  $P1_i$  та  $P2_i$  - набір подій.

При цьому обробка інформаційних даних потоків можна охарактеризувати кількістю операцій  $f_{Pi}(n)$ , щільності їх поділу на певному інтервалі  $f_{Pi}(t)$  і тривалості обробки  $f_{Pi}(T)$ .

У задачі розпізнавання такі інформаційні потоки мають місце у процесах затримки і обчисленні ефективності компонент системи, а дії - операції над даними. Таким чином, можна здійснити оцінку характеристик ефективності інформаційної системи.

Імітаційне моделювання дозволяє визначити характеристики черги і втрати ефективності під час управління інформаційними потоками вибору заявок, заданих параметрів обслуговування та інших. У системі управління та підтримки дозволяє оцінити взаємодію між компонентами та пропускну здатність системи. Проте є обмеження отримання характеристик в реальних умовах через складність умов проведення експерименту.

Моделі модуляційних або пуассонівських потоків часто використовують для обчислення характеристик якості обслуговування з точки зору опису вимог у інформаційних потоках.

Показником ефективності функціонування системи можна використати функцію втрат, використовуючи наступне співвідношення:

$$C^g = \sum_{i=0}^N \alpha_i \lambda_i \bar{t}_i^{(B)}, \quad (4.1)$$

де,  $\alpha_i$  штраф за одиницю часу очікування,  $\lambda_i$  - середня тривалість очікування заявки і-го типу,  $\bar{t}_i$ - інтенсивність і-го типу заявки,  $v$  - параметри характеристики дисципліни обслуговування.

Інтенсивність інформаційних потоків відображає обсяг, швидкість та частоту передачі інформації в певному контексті. У праці [131] визначені основні параметри: довжина маршрутів, час комутації пакета, пропускну здатність каналів зв'язку (бітрейт), довжина пакета. Ефективні параметри інформаційних

потоків - це характеристики, які визначають якість і ефективність передачі, обробки, зберігання та використання інформації в системі і включати такі аспекти [131]:

- *Частоту помилок (error rate)* можна використовувати як один з показників для оцінювання якості функціонування систем на основі бездротових мереж. Схожим способом надійність системи протестовано на основі тестування ефективності.
- *Пропускна здатність (throughput)* є максимальним обсягом інформації, який може бути переданий протягом певного періоду часу. Вищу пропускну здатність можна забезпечити завдяки меншій кількості переданих даних за допомогою запропонованої оптимізації перетворення кількості пікселів із зображення у порівнянні із надсиланням цілісного зображення.
- *Затримка (latency)* середній час передачі або отримання інформації.

Вищезгадані параметри можна відстежувати за допомогою інструменту для візуалізації та аналізу даних Kibana, який часто використовується разом з Elasticsearch для візуалізації даних, що зберігаються у форматі JSON та інших форматах. Метричних візуалізацій дозволяють візуалізувати дані та оцінити інтенсивність інформаційних потоків, такі як використання ресурсів, час відповіді, швидкість транзакцій тощо. Основні результати використаної потужності центрального процесора (CPU), нормалізоване значення завантаження ресурсу, оперативної пам'яті та дискового простору наведені на Рис. 4.2.

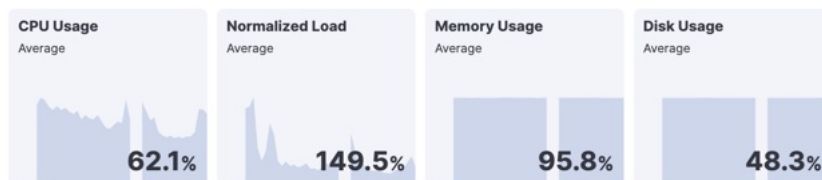


Рис. 4.2. Основні показники ефективності Elasticsearch отримані у середовищі Kibana

Динаміка зміни відповідних навантаження показана на Рис. 4.3. З результатів статистичних звітів Elasticsearch [132] більш детально відображаються показники ефективності різних операцій (зокрема, додавання групованих даних у індекси сховища даних отримано результати про кількість виконаних операцій, загальний та середній час операції та розмір в байтах)

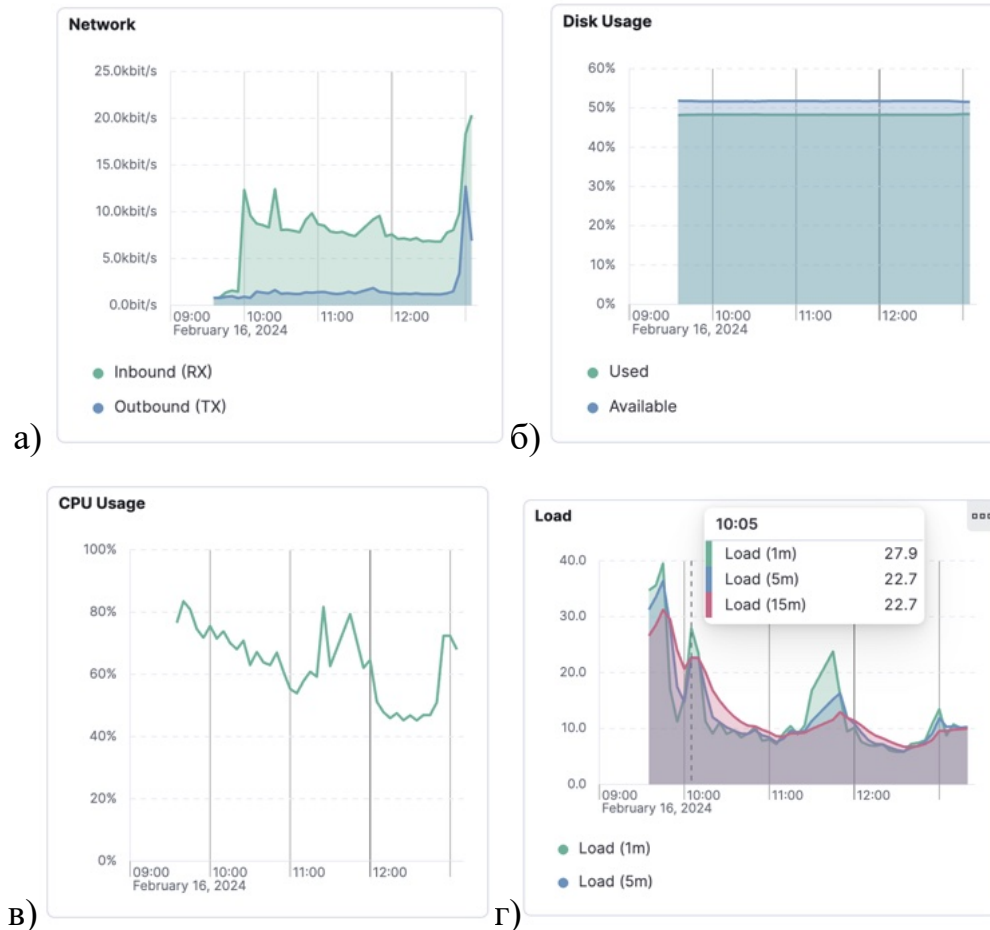


Рис. 4.3. Динаміка зміни основних показників ефективності системи з врахуванням інтенсивності інформаційних потоків а) вхідний (inbound) та вихідний (outbound) трафік, б) завантаження дискового простору, в) використання центрального процесора, г) середнє навантаження системи протягом останньої хвилини, 5 хвилин та 15 хвилин відповідно

Результати інтенсивності інформаційних потоків (Рис. 4.3) показують оптимальність обраного підходу для аналізу даних про дефекти на поверхнях ЕІК та відсутність критичного навантаження. Деякі аспекти оцінювання затримки,

пропускної здатності та частоти помилок апробовані у дослідженні ефективності програмного інтерфейсу [11].

Загальний алгоритм оцінювання стану кіберфізичної системи “МПТ-СКЗ” може виконуватися на основі ефективних параметрів інформаційних потоків та діагностичних показників може складатися з кількох кроків:

1. Визначення найважливіших частин системи та залежностей компонентів на основі функціонального моделювання.
2. Ідентифікувати обмеження програмних частин системи з аналізом факторів або причин низької продуктивності, ефективності чи якості.
3. Обчислення діагностичних показників.

Визначення найбільш важливих частин системи виконане на основі аналізу функціональної моделі кіберфізичної системи “МПТ-СКЗ”. Проведення такого аналізу здійснено на основі відомого засобу створення функціональних моделей об'єкта в конкретній предметній області - технології графічних нотацій SADT (Structured Analysis and Design Technique – технологія структурного аналізу та проектування) [132]. Використовуючи SADT можна здійснювати поетапну ієрархічну декомпозицію системи. Для візуального моделювання існують методології моделювання функцій IDEF0 (Icam DEFinition) та моделювання зв'язків у базі даних IDEF1X, які призначена для розробки і опису моделей даних як показано у праці [133].

Відповідно до концептуальної схеми розроблено візуальну модель функціонування інформаційних потоків у кіберфізичній системі “МПТ – СКЗ” (Рис.4.4).

Методологія IDEF0 дозволяє оцінити важливість супроводу на певних етапах реалізації системи на основі проектованої кількості встановлених зв'язків між функціями.

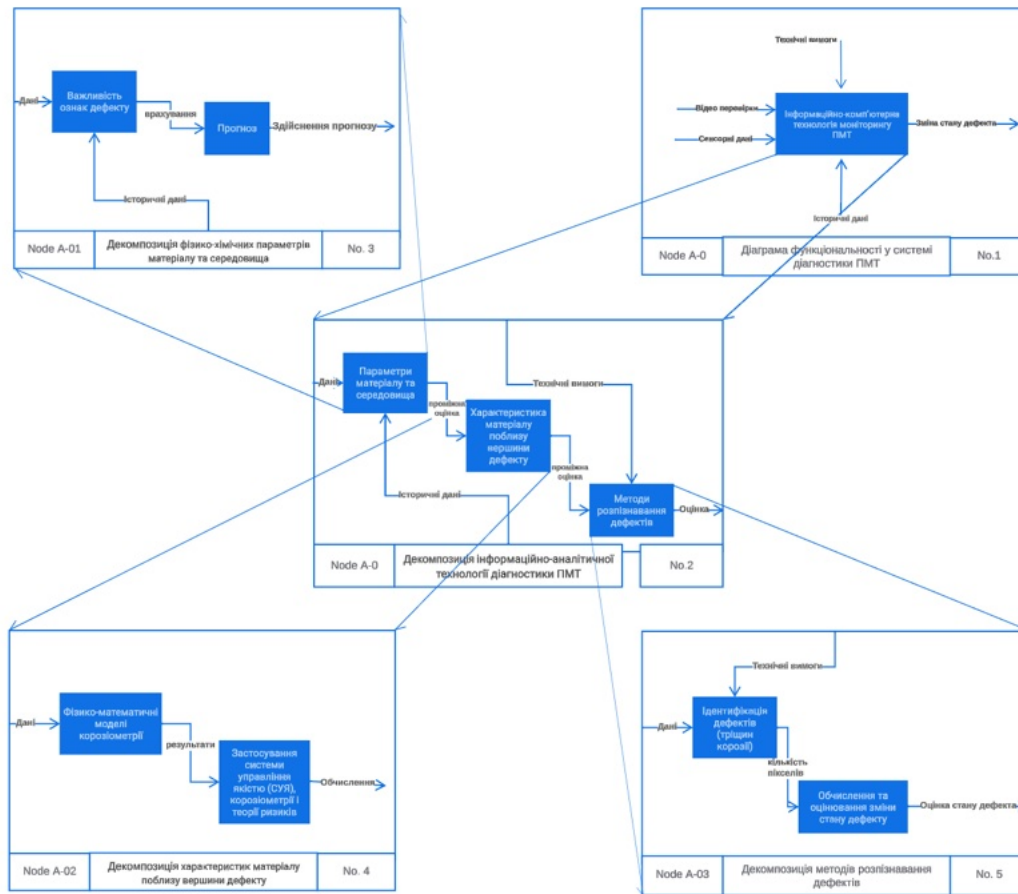


Рис. 4.4. Функціональна діаграма інформаційних потоків

На Рис.4.4 зазначені компоненти можуть бути складовими моделі прийняття рішень в контексті дефектоскопії або аналізу стану дефектів. Перший в цій моделі є етапі визначення параметрів матеріалу та середовища можуть бути важливими для визначення станів дефектів на основі низки факторів впливу. Наступним доповнюючим кроком є формування характеристики матеріалу поблизу вершини дефекту, що дозволяє описати параметри матеріалу в місці дефекту, глибину дефекту, тип матеріалу, його структуру, можливість ідентифікації особливостей поверхні тощо. Валідаційним етапом є методи розпізнавання дефектів для ідентифікації та класифікації дефектів. Відповідні методи використовують технології машинного навчання, обробку зображень та інші техніки для автоматичного визначення дефектів на основі зібраних даних.

Табл. 4.1. Перелік та характеристика граничних зв'язків функціональної моделі інформаційно-комп'ютерної технології МПТ

Назва зв'язку	Тип зв'язку	Кількість процесів або функцій	Моделі прийняття рішення		
			Параметри матеріалу та середовища	Характеристика матеріалу поблизу вершини дефекту	Методи ідентифікації дефектів
Дані	Input	3	1	1	1
Історичні дані	Mechnism	1	1		
Технічні вимоги	Control	1			1
Зміна стану дефекта	Output	3	1	1	1
Ступінь важливості	%	100	37,5	25	37,5

У Табл. 4.1 важливим етапом є методи розпізнавання як валідаційний крок, що дає змогу здійснювати оцінювання зміни стану дефекту та прийняття рішення у технічному обслуговуванні ЕІК.

Ступінь важливості інформаційного потоку в кожному етапі - це оцінка оцінку рівня деталізації, точності та своєчасності інформації, яка надходить на кожному етапі. Це допомагає визначити, які аспекти процесу потребують особливої уваги та контролю з точки зору забезпечення ефективності, надійності та безпеки.

#### **4.2. Розв’язування та дослідження варіантів задачі оптимізації для оцінювання якості кіберфізичної системи “МПТ – СКЗ”**

Запропоновану у роботі модель можна використати для розширення в контексті оцінювання якості кіберфізичної системи металеві підземні трубопроводи - система катодного захисту “МПТ – СКЗ”.

Проблема отримання якісної інформації про стан трубопровідних конструкцій є важливою для технічної діагностики і екологічної безпеки газотранспортних систем (ГТС). Адже під час руйнування підземних магістральних трубопроводів й миттєвого вивільнення енергії виникають механічні пошкодження природного ландшафту та рельєфу, порушення цілісності ґрунтово-рослинного покриву [134]. Під час загоряння газу можливе пошкодження територій радіусом до 550 м від осередку аварії, а фрагменти трубопроводу після вибуху можуть поширюватись до 480 м [134].

У зв’язку з цим для контролю підземних металевих трубопроводів (ПМТ) доцільно використовувати сучасні методи моніторингу у поєднанні з інформаційними технологіями (ІТ) і комп’ютерним моделюванням, а також з урахуванням методів інженерії якості та механіки руйнування твердих тіл.

У роботі [135] розглядається специфіка моделювання якості складних систем з допомогою контролю якості ЕІК.

Для удосконалення процедури діагностування такого типу систем з допомогою засобів інформаційно-вимірюваної техніки вводиться діагностична вага ознак та діагностична цінність обстежень, що сприятиме підвищенню якості досліджень [135].

У праці [136], викладено методологію діагностування станів складних промислових об’єктів, яка може бути використана під час створення засобів та пристроїв керування, зокрема, для підземних металевих трубопроводів (ПМТ).

Розглядаються процеси життєвого циклу (ЖЦ), пов’язані із забезпеченням якості комп’ютерних систем (КС) – процеси верифікації, валідації, тестування,



виміри, забезпечення гарантії якості, керування ризиком і якістю у області досліджень сфери інженерії якості [137].

Визначено напрямки розвитку інженерії якості КС, які розробляються у новій парадигмі програмування – генеруюче програмування (ГП) [137]. Охарактеризовано актуальні проблеми забезпечення якості сімейств систем, що складають базис для побудови окремих КС – членів сімейств – з повторно використовуваних компонентів. Отже, у роботі [137] сформульовано задачі інженерії якості сімейств та підходи до їх розв’язання, зокрема задач верифікації, тестування, моделювання та оцінювання якості КС у новій парадигмі.

Проведено огляд основних технічно-аналітичних перетворень інформації корисних для контролю стану пітингостійкості металоконструкції за участю даних реставрації глибини пітинга [138]. На основі результатів контролю стану пітингостійкості металів представлено практичний результат реставрації глибини пітинга на поверхні металу [138]. Наведено опис інтерфейсу та особливостей роботи із розробленою експертною системою оцінювання ризику руйнування конструктивних елементів теплоенергетичної системи [139]. Під час визначення допустимих розмірів дефектів враховується планований термін експлуатації елементів трубопроводу [139].

Основною процедурою моніторингу ПМТ є контроль корозійних дефектів аналогічно як у роботі [140]. Тому запропоновано розширити поточну реалізацію додатковим рівнем перевірки на основі значень вимірних потенціалів.

Система корозійного моніторингу трубопроводів (СКМТ) призначена для визначення місць пошкодження захисного покриття, вимірювання потенціалів (корозійних, поляризаційних, захисних, поперечного і поздовжнього градієнтів) трубопроводу, корозійної активності ґрунту (швидкості корозії металу трубопроводу на різних рівнях його залягання) і швидкості залишкової корозії металу трубопроводу в дефектах захисного покриття [140].

Критерієм пошкодження захисного покриття трубопроводу є поляризаційний потенціал  $E_P$  та його відхилення  $D_{E_P}$ , що пов'язано з катодним захистом і обумовлено відповідними нормативними документами [140]. Вимірювання  $E_P$  проводиться при відключеній системі КЗ на контрольованій ділянці труби методом виносу електрода порівняння, що не поляризується.

У програмному забезпеченні системи моніторингу (СМ) і системи катодного захисту ПМТ закладемо програму автоматичного розрахунку корозійної активності ґрунту (тобто густину струму корозії  $i_K$ ) за формулою [139]:

$$i_K = \frac{2B\Delta I}{S\Delta E}, \text{ мм/рік, (1мм/рік} = 0,8616 \text{ A/m}^2) \quad (4.2)$$

де  $B$  – постійна методу (для трубних сталей має значення 332);  $S$  – площа поверхні одного електрода сенсора для оцінки швидкості корозії (ШК),  $\Delta I$  – виміряне значення електричного струму, А;  $\Delta E$  – різниця потенціалів, що задається між робочими електродами сенсора ШК (складає 10 mV).

Після визначення корозійної активності ґрунту  $i_K$  (густини струму) можна розрахувати швидкість  $i_D$  (густини струму) залишкової корозії металу трубопроводу у дефекті захисного покриття за формулою [140]:

$$i_D = i_K \cdot 10^{V_b}; V_b = (E_{cor} - E_P)/b_a, \quad (4.3)$$

де  $E_{cor}$  – потенціал корозії металу трубопроводу, V;  $E_P$  – поляризаційний потенціал в точці вимірювання, V;  $b_a$  – нахил анодної поляризаційної кривої у рівнянні Тафеля, V.

Рівняння Тафеля зв'язує перенапруження електродної реакції  $h$  розчинення металу (сталі) з густиною корозійного (анодного) струму  $i_A$  [140]:

$$\eta = a + b \cdot \lg \frac{i_A}{i_N}, a = \frac{2,3RT}{(1-\alpha)nF} \lg \frac{i_A}{i_N}, b = \frac{2,3RT}{(1-\alpha)nF}, i_A = i_0 \cdot \exp \left[ \frac{(1-\alpha)nF}{RT} \eta \right]. \quad (4.4)$$

Тут  $i_N = 1 \text{ A/m}^2$  – нормувальний множник, який забезпечує безрозмірне відношення під знаком логарифма;  $n$  – кількість електронів, які беруть участь в електродній реакції розчинення металу;  $T$  – температура;  $R$  – універсальна газова

стала;  $F = 96485$  Кл/моль – стала Фарадея;  $a$  – коефіцієнт переносу;  $i_0$  = струм обміну (який відповідає умові  $|i_A| = |i_C| = i_0$ ;  $i_C$  – катодний струм).

Найбільш інтенсивно процес корозії проходить у вершині тріщини і тому для більш детального аналізу анодного розчинення металу доцільно врахувати співвідношення, яке уведено у монографії Кеше [141] і узагальнено у праці [142]:

$$i_A = \frac{\alpha \cdot \chi \cdot \Delta \psi_{ak}}{\delta \cdot \ln(c/\delta)} \cdot \left( 1 + \beta_W \cdot \left( \frac{W_{PL} - W_{PL0}}{W_{PL0}} \right) S \right), \quad (4.5)$$

де  $a$  – кут у вершині поверхневого дефекту (тріщини, пітинга);  $c$  – електропровідність електроліту (зокрема, ґрунтового);  $\Delta \psi_{ak}$  – омічна зміна електричного потенціалу між анодною (А) та катодною (К) ділянками;  $c$ ,  $d$  – ефективна глибина і розкриття тріщини відповідно;  $b_W$ ,  $S$  – емпіричні константи;  $W_{PL}$  – поверхнева енергія пластичного деформування (ПЕПД) за умови напруженого стану  $s/s_T = 1 \div 1,3$  в діапазоні пластичних деформацій;  $s$  – механічне напруження;  $s_T$  – межа текучості;  $W_{PL0}$  – ПЕПД за умови відсутності зовнішнього механічного навантаження. Співвідношення (4) записано для вершини тріщини, тобто анода А. Бокові поверхні дефекта покриті слабо провідними окислами (пасивними плівками) і є катодом К.

Співвідношення (1)-(4) є основою удосконаленої моделі корозіометрії (рисунок 4.5) і відповідної інформаційної технології порівняно з [142].

Інформаційна технологія (рисунок) системи моніторингу ПМТ поєднує 4 типи моделей:

- математична модель корозіометрії з урахуванням контролю якості підземних трубопроводів в контексті виявлення поверхневих дефектів (тріщин, пітингів) [134, 135, 140, 142, 143];

- моделювання процесів з погляду системи управління якістю (СУЯ) [161];

- моделювання процесів з погляду організаційних структур (персонал, інженерія якості, забезпечення експлуатаційної безпеки технологічного процесу ГТС) [136, 138, 145];

– моделі дефектів і методики їх розпізнавання, а також критерії міцності відповідних елементів конструкцій з дефектами (тріщинами, пітингами) [10, 14, 147].



Рис. 4.5. Структурна схема інформаційно-комп'ютерної технології з урахуванням корозійних процесів поблизу вершини тріщини, СУЯ, ризиків

Уведемо в розгляд інтегральний показник ефективності функціонування системи моніторингу ПМТ аналогічно як у [145, с. 136] з доповненням параметрами системи управління якістю (СУЯ) [135]:

$$E_F = f(F(R), F(D), F(Q), F(I_t, P_t)) \Rightarrow opt. \quad (4.6)$$

Тут  $F(R)$  – ефективність використання ресурсів з погляду моделі корозіометрії;  $F(D)$  – ефективність діяльності структурних підрозділів ГТС;  $F(Q)$  – ефективність в результаті функціонування СУЯ;  $F(I_t, P_t)$  – функція результативності, а також ефективності діяльності працівників (персоналу);  $I_t$  – індекс творчого потенціалу, кваліфікації та лояльності працівників;  $P_t$  – модель вибору поведінки працівників (персоналу), включаючи якісні і кількісні

параметри  $P_{i_i}$ . ( $i = 1, 2, \dots, n_p$ ;  $n_p$  – загальна кількість параметрів відповідної моделі).

Для оптимізації методики розпізнавання дефектів, а також інформаційних потоків  $P_k(X_i)$  в системі моніторингу ПМТ і покращення системи протикорозійного захисту підземних трубопроводів використаємо аналогічно як у [146] функціонал якості з урахуванням оберненого зв'язку:

$$J(P_k(X_i), FB(X_i)) = \int_{t_0}^{t_k} f(\bar{y}, \bar{u}, \bar{\zeta}) dt \Rightarrow opt, \quad (4.7)$$

де  $\bar{y}$  – вектор заданих впливів ( $y_j(t)$  – компоненти вектора,  $j = 1, 2, \dots, n$ );  $\bar{u}$  – вектор керувань;  $\bar{\zeta}$  – вектор невизначених збурень;  $[t_0, t_k]$  – інтервал часу, в якому розглядається процес (формування оптимальних значень інформаційних та фінансових потоків  $P_k(X_i)$ ;  $k=1, 2, \dots, m$ ;  $i=1, 2, \dots, m_p$ ;  $m$  – загальне число інформаційних та фінансових потоків, які мають відношення до даної системи (ГТС);  $m_p$  – загальне число параметрів);  $\bar{\zeta}$  – функція, що відображає показник якості;  $FB(X_i)$  – функція, яка характеризує обернений зв'язок (*Feed-back*) між потоками  $P_k$  і оточенням ПМТ з урахуванням думок експертів. Тут символ *opt* відповідає умові оптимальності функціоналу.

Доцільно уважати, що обернений зв'язок  $FB(X_i)$  пов'язаний з ризиками та параметрами, які характеризують якість, надійність і точність параметрів, які будуть отримані на основі запропонованих в даній праці підходів. Для оптимізації ризиків врахуємо такі фактори [146, 147]: якості та надійності, інформаційної ємності і фактори ризику, пов'язані з програмним забезпеченням.

За сучасних темпів розвитку індустрії програмного забезпечення є важливим формування проблемної області (ПрО), підвищення якості програмних систем (ПС), побудова ПС у новій парадигмі генеруючого програмування (ГП) з множини різнотипних повторно використовуваних програмних компонентів, а також постійне накопичення, узагальнення та використання знань стосовно ПрО та перспектив їх еволюції [137]. Інженерію ПрО ототожнюють з інженерією сімейств ПС [137].

Пошук напрямків розв'язання проблем інженерії якості стосується системи моніторингу ПМТ (рис.4.1). Зокрема, це такі проблеми [137]:

- 1) моделювання якості сімейства ПС та оцінювання якості генерованих артефактів на кожній стадії інженерії ПрО та інженерії застосувань;
- 2) вибір компонентів, які щонайкраще відповідають вимогам до цільової ПС;
- 3) верифікація активів ПрО. Проблема полягає у відсутності методів перевірки відповідності вибраних (розроблених) активів потребам ПрО, зокрема доведення коректності моделей, компонентів, архітектурних композицій компонентів;
- 4) тестування компонентів нової генерації. Проблема полягає в необхідності пристосовування існуючих методів онлайнового (оперативного) тестування вибраних компонентів для генерації ПС, зокрема, Web-компонентів;
- 5) розроблення інструментів тестування. Видається, що інструменти тестування компонентів ПС та їх композицій, як специфічний вид ПС у ПрО тестування, самі могли б утворювати сімейства, що полегшило б їх пристосування до програмної архітектури середовища IDE, в якому розробляються цільові ПС;
- 6) сертифікація компонентів – затвердження їх відповідності загальноприйнятим стандартам, а також адекватності заданій множині вимог.

Проблема полягає у тому, що початковий код вибраних компонентів (COTS-продуктів або продуктів з Інтернет) недоступний для аналізу і верифікації відомими методами. Необхідно комбінувати різні стратегії сертифікаційного тестування, включаючи тестування не тільки розробниками, але і постачальниками, і споживачами цих компонентів;

7) визначення оперативної програмної архітектури, що не буде протирічити вимогам до якості ПС. Проблема полягає у тому, що програмна архітектура ПС, яка вибудовується з архітектури сімейства шляхом композиції, має формуватися з урахуванням не лише функціональних вимог до ПС, але й вимог до якості.

8) керування ризиком в ході всього життєвого циклу (ЖЦ) генерації СУЯ. Проблема полягає в необхідності аналізу компромісів під час вибору тієї або

іншої архітектурної композиції, стратегії генерації ПС, тестування та оцінювання якості. Відповідно розв'язання задач пов'язано з визначенням таксономії ризиків і регулярним оцінюванням ризиків на кожному етапі генерації ПС.

### 4.3. Статистична обробка даних для оцінювання критичного стану функціонування інженерного об'єкта з дефектами

Ймовірність виявлення дефекту обчислюється враховуючи низку властивостей функції розподілу [67]:

- кількість дефектів буде меншою, якщо розміри дефектів  $X$  є більшими і суттєво відрізняються від середнього значення  $X$ .
- дефекти не можуть бути нескінченно великих розмірів, тому вважаємо що  $\lim_{x \rightarrow \infty} f(x) = 0$ .
- Відносно значення в точці  $X$  функція розподілу повинна бути парною:  $f(\bar{X} - X) = f(\bar{X} + X)$ .
- найбільша кількість дефектів повинна мати розміри  $X_i$ , близькі до середньої величини  $X$ , причому функція розподілу буде найвищою в цій області  $f(\bar{X}) = f_{max}$ .

Однією з функцій, що задовольняє всім цим властивостям, є функція Гауса:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right), \quad (4.8)$$

Ймовірність виявлення дефекту дорівнює значенню інтеграла від заданого розміру дефекту  $x_0$  до  $x$  (з розміром  $x > x_0$ ) можна представити у вигляді різниці двох інтегралів:

$$P(x > x_0) = \int_{\bar{x}}^{\infty} f(x)dx - \int_{\bar{x}}^{x_0} f(x)dx, \quad (4.9)$$

Підставивши в цей вираз функцію Гауса отримаємо

$$P(x > x_0) = \frac{1}{\sqrt{2\pi}\sigma} \left( \int_{\bar{x}}^{\infty} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right)dx - \int_{\bar{x}}^{x_0} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right)dx \right), \quad (4.10)$$

Інтеграла похибок  $\Phi(y)$  максимально допустимого розміру дефекту  $x_0$  може бути обчисленим на основі математичного очікування  $x$  та дисперсію  $2\sigma$

(або  $\sigma$ ). Ймовірність знаходження дефектів за розмірами більшими  $x_0$  визначається наступним чином:

$$P(x > x_0) = \frac{1}{2}(1 - \Phi(y)), z = \frac{x_0 - \bar{x}}{\sigma}, \quad (4.11)$$

де  $\Phi(y) = \sqrt{\frac{2}{\pi}} \int_0^y \exp(-\frac{t^2}{2}) dt$  - інтервал похибки для  $\int_0^\infty \exp(-\frac{t^2}{2}) dt$ .

При цьому у галузі дефектоскопії прилади, які ідентифікують дефекти, повинні бути налаштовані на знаходження 99 зі 100 можливих дефектів, що означає  $P(x > x_0) = 0,99$  [67].

У системах контролю визначають придатними до експлуатації тільки ЕК без неприпустимих (критичного стану) дефектів (з параметрами  $x \geq x_0$ ). Невиконання цієї умови спричиняє відмову системи, яка передбачає неправильне віднесення (недобраковку) або помилкове бракування (перебраковку) певного об'єкта. Ймовірності недобраковки ( $F$ ) та перебраковки ( $\beta$ ) часто використовують для оцінки працездатності системи контролю, її надійності, достовірності, ефективності. Це означає, що показники надійності характеризують здатність досліджуваних об'єктів до продовження функціонування за умови дотримання конкретних умов експлуатації та технічного обслуговування. У системах контролю, одним з кількісних показників надійності системи є імовірність безвідмовної роботи, яка визначається формулою аналогічно як у [67].

$$G = 1 - (F + \beta). \quad (4.12)$$

Наслідки неправильного виявлення та невірної класифікації дефектів відрізняються, тому для оптимізації системи контролю використовуються різні критерії. Зазвичай найбільш оптимальною вважається система, яка мінімізує середній ризик, який оцінюється за допомогою функції  $R$ :

$$R = c_1\beta + c_2F, \quad (4.13)$$

де  $c_1$  і  $c_2$  - вартість перебраковки і недобраковки.



Найоптимальніші значення зазвичай спостерігаються при рівній вартості  $c_1$  і  $c_2$ , тоді система може забезпечити мінімум середнього ризику, і максимум імовірнісного показника надійності  $G$ , обчисленого за формулою (1.10). Значення  $R$  зазвичай встановлюється відповідно до бракувального рівня. Такий чином це влаштовує постачальника послуг, який зацікавлений в мінімальному значенні показника перебраковки, і замовника технічного обслуговування, який зацікавлений в мінімальному значенні показника недобраковки. У разі більш ретельного контролю, наслідки виходу з експлуатації частин ЕІК є більша за витрати на перебраковку, обирають уточнений показник з ймовірністю відсутності недобраковки  $G_0 = 1 - F$ .

Під час оптимізації прийняття рішення може відбуватися фіксація більш дрібних дефектів з подальшим додатковим аналізом зі зниженням границь допустимих зчитаних значень. При такому режимі роботи у підвищеній чутливості системи контролю системи на рівні  $X \leq X_0$  зростає перебраковка. Тому, для зменшення переробки повторно проводиться уточнення іншими засобами попередньо зафіксованих дефектів. Таким чином, вдається досягти досить високого рівня показника надійності  $G_0 = 0,997$  і відповідно підвищення надійності системи контролю зі збільшенням додаткових витрат (до 5%).

В теорії надійності часто використовують Критерій Вейбулла (або розподіл Вейбулла), особливо під час аналізу даних за допомогою статистичних методів контролю якості та ідентифікації дефектів у виробництві. В основному, відповідний метод оцінювання надійності застосовується для аналізу часу до виникнення подій, таких як відмови обладнання або дефекти виробів. Цей розподіл дозволяє моделювати часовий інтервал між виникненням подій і можуть мати один, два або три параметри [46].

Функція розподілу Вейбулла для двох параметрів має вигляд [46]:

$$f(x; \lambda, k) = \left\{ \frac{k}{\lambda} \cdot \left(\frac{x}{\lambda}\right)^{k-1} \cdot e^{-\left(\frac{x}{\lambda}\right)^k} \text{ для } x \geq 0, \text{ і } 0 \text{ для } x < 0 \right\} \quad (4.12)$$

де  $x$  - час до виникнення події (дефекту),  $\lambda$  - параметр масштабу, що визначає, як швидко виникають події,  $k$  - параметр форми, який визначає форму розподілу (при  $k=1$  отримуємо експоненційний розподіл).

Критерій Вейбулла в контексті ідентифікації та аналізу дефектів використовується для моделювання часу до виникнення дефектів чи відмов у виробництві, виявлення аномальних відмов чи дефектів, які не підпадають під типовий розподіл та визначення оптимального графіку технічного обслуговування та запобігти непередбаченим відмовам.

Метрикою якості відбору параметрів критерію Вейбулла є логарифм правдоподібності, який стосується процесу визначення найкращого розподілу даних з урахуванням залежності в даних. У порівнянні із задачами класифікації аналогічним показником є точність або у регресії значення  $R$  у квадраті. Зважаючи на обмеження показника логарифма правдоподібності у випадку багатьох параметрів або невеликої кількості вибірок, використовують інформаційний критерій Akaike (AIC) та байєсівський інформаційний критерій для невеликих та великих кількостях вибірок відповідно. Показник Андерсона-Дарлінга (AD) показує розподілення даних відповідно до припущення Вейбулла [148].

```

Results from Fit_Weibull_2P (95% CI):
Analysis method: Maximum Likelihood Estimation (MLE)
Optimizer: TNC
Failures / Right censored: 10/0 (0% right censored)

Parameter Point Estimate Standard Error Lower CI Upper CI
Alpha      1163.46          179.179    860.322   1573.4
Beta       2.17087           0.541846   1.33099   3.54073

Goodness of fit Value
Log-likelihood -75.8891
AICc 157.492
BIC 156.383
AD 1.47674

```

Рис. 4.6. Результати виконання Fit\_Weibull\_2P

Для реалізації таких перевірок використано метод Fit\_Weibull\_2P (рис. 4.6) з бібліотеки Reliability у мові програмування Python, що показує залежність введеного параметру форми ( $k=1.7$ ) і виводу масштаб 1163.46 та форму функції

ймовірності 2.17. Зазвичай зі збільшенням масштабу швидше відбувається збій, а рівень відмов зростає з часом при формі більше 1.

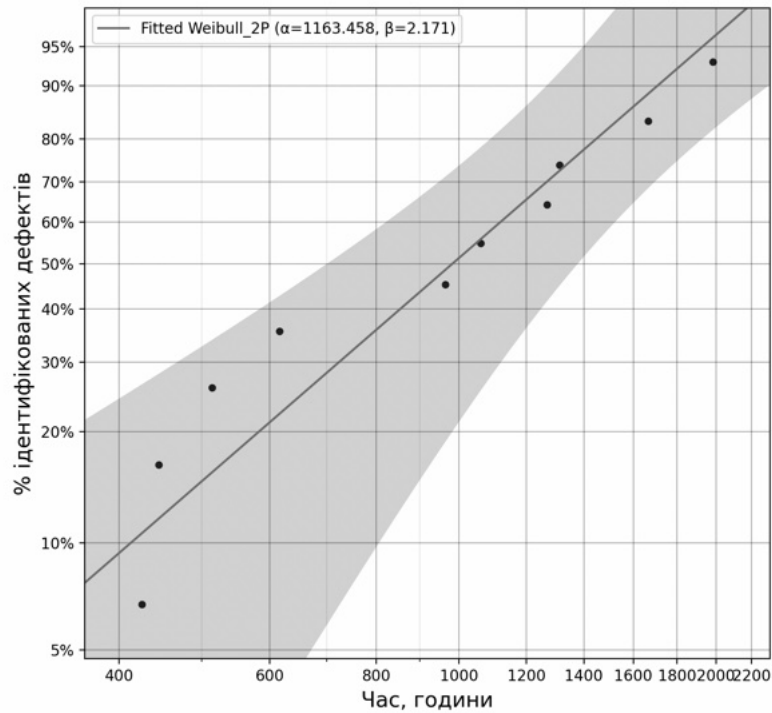


Рис. 4.7. Імовірнісна схема Вейбулла

Серед відомих методів довірчих меж використовують межі Фішера, межі відношення правдоподібності, межі бутстрапа, бета-біноміальні межі, опорні межі Монте-Карло та інші. Параметр форми  $\beta$  представляє нахил лінії Вейбулла та описує режим відмови, а параметр масштабу  $\eta$  визначається як значення осі  $x$  для ненадійності 63,2 %.

Як вхідні дані для побудови функції розподілу використано час до ідентифікації перших дефектів на основі кількох відеозаписів. Період від початку аналізу до знайденої першої кількості помилок можна обчислити за час ідентифікації дефектів на поверхнях і можна використати для того, щоб показати функцію щільності на Рис. 4.8. Для візуалізації використано метод `weibull_pdf` з бібліотеки `predictr` у мові програмування Python.

Розподіл Вейбулла є статистичним розподілом, який використовується для моделювання часу між настанням подій, таких як час між прибуттям пакетів

даних в мережі. Наприклад, для наборів даних [3, 3, 3, 3, 3, 3, 4, 4, 9], [3, 3, 5, 6, 6, 4, 9] та [5, 6, 6, 6, 7, 9] результати показані на рис. 4.8.

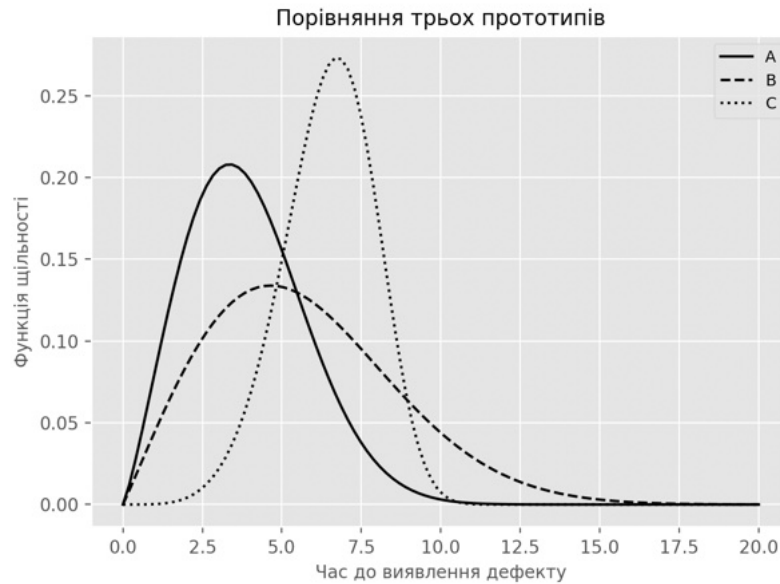


Рис. 4.8. Функція щільності трьох прототипів даних у схемі Вейбулла

Проведений аналіз надійності може бути корисний для прийняття рішення у визначенні пріоритетів обмеженого фінансування ремонту, забезпечуючи комплексну кількісну оцінку продуктивності за допомогою запропонованої автоматизованої системи.

#### 4.4. Метод обчислення ефективних параметрів діагностики, які характеризують покращення якості кіберфізичної системи “МПТ – СКЗ”

Оцінювання пошкоджень, яке виконано згідно з формулою (1.1), можна розширити із застосуванням низки додаткових параметрів обчислення ефективності виконаної діагностики на основі кількісних даних про ідентифіковані дефекти на поверхні. Відповідне покращення оцінювання якості систем запропоновано доповнити кількома простими параметрами:

- Загальна кількість проведених діагностик  $D_{\text{загальних}}$  (тестових випадків) відповідно до запропонованої моделі відповідає кількості аналізів потоку зображень. У проведеному дослідженні виконано аналіз 7-и об’єктів (зокрема, елементи труб, мостових конструкцій та доріг).

- Загальна кількість ідентифікованих дефектів  $N_{загалом}$  під час діагностики, що відповідає загальній кількості записів (docs) про ідентифіковані дефекти на поверхнях в Elasticsearch як показано на Рис. 4.9. (корозійних - 2003, типу тріщин - 2998)

<input type="checkbox"/> Name	Health	Status	Primaries	Replicas	Docs count	Storage size
<input type="checkbox"/> corrosion_index	● green	open	1	0	2 003	108.3kb
<input type="checkbox"/> crack_index	● green	open	1	0	2 998	159.58kb

Рис. 4.9. Перегляд основних параметрів у процесі управління індексами Elasticsearch

- Середня кількість ідентифікованих дефектів  $N_{середня}$  під час діагностики, що може бути поділена по часу діагностики до загальної кількості відповідно.

Більш уточнені значення ефективних параметрів оцінювання виконаної діагностики (D) дефектів на поверхнях може вимагати більшої кількості параметрів, які аналогічними з теорії тестування програмних продуктів [149] та доповнюючими :

- Відсоткова кількість ідентифікованих дефектів критичного стану на основі збережених даних про стан (індекс device\_index) в Elasticsearch.

$$P_{критичних} = \frac{N_{критичних}}{N_{загалом}} \cdot 100$$

- Протилежним до попереднього є відсоткова кількість ідентифікованих дефектів некритичного стану.

$$P_{некритичних} = \frac{N_{некритичних}}{N_{загалом}} \cdot 100$$

- Кількість ідентифікованих дефектів за одну проведену діагностику  $N_{дефектів за одну діагностику}$ , що визначається за фільтруванням по часу із збережених даних.

$$N_{дефектів за одну діагностику} = \frac{N_{загальних}}{D_{загалом}}$$

Оскільки процес є автоматизований, то методи оцінювання можна розширити враховуючи практику використання автоматизованого тестування програмного забезпечення, а саме:

- Середній час ідентифікації дефектів або швидкість пошуку дефектів  $t_{\text{середній}}$  за загальний час виконання діагностики (в годинах)  $t_{\text{загалом}}$

$$t_{\text{середній}} = \frac{N_{\text{загалом}}}{t_{\text{загалом}}}$$

- Ефективності діагностики якості  $E_D$  визначається як кількість аварій до загальної кількості знайдених дефектів.

$$E_D = \frac{N_{\text{аварій}}}{N_{\text{загалом}}}$$

- Ефективність ідентифікації дефектів  $E_I$  на основі їх співвідношення з ідентифікованих та прогнозованих значень

$$E_I = \frac{N_{\text{ідентифікованих}}}{N_{\text{прогнозованих}}} \cdot 100$$

- Індекс серйозності дефекту (Defect severity index - DSI) є сумою кількості ідентифікованих дефектів за критерієм та індексів серйозності (SI) (згідно з таблицею ймовірності аварій до стану дефекту).

$$DSI = \frac{\sum_{i=0}^n (N_{\text{дефектів за критерієм}} \cdot SI)}{N_{\text{загалом}}}$$

де  $N_{\text{дефектів за критерієм}}$  – кількість дефектів різних станів,  $SI$  – коефіцієнт серйозності дефекту. Описані показники та запропонований метод обчислення ефективних параметрів діагностики, які характеризують покращення якості кіберфізичної системи "МПТ – СКЗ" дає змогу визначити показники якості системи, обчислити ефективних параметрів діагностики, оцінити результатів та покращити функціонування складних систем.

## Висновки до розділу 4

У даному розділі дисертаційної роботи побудовано структурну схему концептуальної моделі інформаційно-комп'ютерної технології захисту підземних трубопроводів з урахуванням методик розпізнавання дефектів, критеріїв якості та міцності елементів конструкцій, ризиків, а також корозійних процесів поблизу вершин тріщин та пітингів.

Оцінювання ефективності інформаційних потоків, які відповідають розглянутим процесам, здійснено з урахуванням показників урівноваження робочого навантаження, створення резервних копій і відновлення даних у сховищах даних. Використано інструмент Kibana для візуалізації динаміки ефективності системи на основі показників таких ресурсів як навантаження на графічний процесор, розмір дискового простору, зміна вхідного та вихідного трафіку мережі. Крім цього, розглянуто методи статистичної обробки даних для оцінювання критичного стану інженерного об'єкта з дефектами у процесах функціонування.

Метою статистичного моделювання процесів є

- формування та оптимізація інформаційних потоків в системі моніторингу ПМТ;
- покращення системи протикорозійного захисту підземних трубопроводів з використанням функціоналу якості, а також з урахуванням оберненого зв'язку і принципів інженерії якості.

Отримані на основі статистичного моделювання результати, викладені у цьому розділі, доповнюють результати, які отримано з допомогою моделі інформаційно-комп'ютерної технології для інтелектуального аналізу змін станів дефектів та можуть бути використані для прийняття рішень стосовно оптимізації нафтогазових транспортних систем у період нестабільних та кризових ситуацій. Результати, описані у цьому розділі, містяться у наступних працях дисертанта [1, 11, 16, 17].

## ВИСНОВКИ

У дисертаційній роботі розроблено нові підходи до прогнозування, ідентифікації та оцінювання пошкоджень у елементах інженерних конструкцій на прикладі зміни стану дефектів на поверхнях цих об'єктів. Відповідні підходи ґрунтуються на поєднанні методів машинного і глибокого навчання, статистичної обробки даних про зовнішні впливи, інтенсивності інформаційних потоків із застосуванням технічних засобів відбору, опрацювання та аналізу даних. На основі низки проведених досліджень отримано наступні основні результати:

1. Розроблено комплексний підхід для оцінювання поверхневих дефектів елементів інженерних конструкцій (ЕІК) на основі методів інтелектуального аналізу даних з використанням технічних засобів та методів комп'ютерного зору, що дозволяють вдосконалити якість та надійність системи підтримки прийняття рішень у комплексних інформаційно-аналітичних системах контролю зображень дефектів;
2. Розроблено та апробовано програмне забезпечення для опрацювання, збереження та аналізу даних про ідентифікацію та оцінювання аспектів динаміки змін станів дефектів на поверхнях з використанням мови Python 3, технічного засобу JetRacer Kit на основі Nvidia Jetson Nano, програмного каркасу YOLO/CNN, бібліотеки sklearn в межах архітектури REST згідно підходів IoT, які забезпечують методи та засоби ефективного діагностування поверхні елементів інженерних конструкцій в режимі реального часу.
3. Запропоновано удосконалений варіант аналізу та оцінювання ресурсу інженерного об'єкту, тобто трубопроводу (конструкції) і системи катодного захисту (СКЗ) з урахуванням удосконаленої моделі, яка дозволяє більш точно проводити ідентифікацію та прогноз зміни станів дефектів на поверхнях.



4. Подальший розвиток отримав підхід оптимізації послідовності процедур прийняття рішень про критичний стан дефектів на поверхні конструкції із урахуванням впливу середовища, а також інформації про історичні дані. На основі цього підходу побудовано процес формування послідовності оптимальних рішень, які характеризують поведінку дефектів на поверхні об'єктів і мають відношення до складних інформаційних середовищ.
5. Удосконалено функціональну структуру інформаційних потоків і завдяки цьому проведено оцінювання оптимальних значень кількості кадрів в секунду для ефективного опрацювання зображень ЕІК з поверхневими дефектами засобами Nvidia Jetson Nano та Raspberry Pi. Отримані відповідні результати дозволили раціонально проводити подальший аналіз зображень дефектів та визначати достовірні значення ресурсу елементів конструкцій на цій основі;
6. Запропоновано спосіб представлення даних щодо передачі зображень поверхневих дефектів ЕІК у бази даних для їх ідентифікації і контролю засобами IoT, який базується на групуванні виділених пікселів, що характеризують взаємодію даних з моменту відбору інформації про показники середовища, в якому функціонує об'єкт. Відповідний підхід дозволив не лише моніторити стан об'єкта, але і забезпечив автоматизований процес аналізу та прогнозу змін у дефектах, що дало можливість ефективно управляти системою катодного захисту, приймати своєчасні рішення та формулювати корисні рекомендації, а також перевіряти результати прогнозу щодо змін стану дефектів на поверхні елементів конструкцій.
7. З урахуванням підходів інженерії ознак розроблено метод оцінювання змін стану дефектів. На цій основі ідентифіковано характерні особливості поведінки основних інформативних

параметрів, які відповідають зображенням дефектів, та проаналізовано методами машинного навчання характерні особливості поведінки інформаційних потоків, які дозволяють передбачати зміни стану дефектів, вдосконалити процедуру їх моніторингу та визначити елементи стратегії попередження можливих ризиків.

8. Проаналізовано та систематизовано наявні методи аналізу даних та системи діагностування ЕІК, що дало можливість виділити основні інформативні параметри та з їх допомогою побудувати модель прогнозування змін стану поверхневих дефектів на базі регресійних методів. В результаті обчислювального експерименту встановлено, що запропонована модель дозволяє підвищувати якість автоматизації роботи діагностичних систем та забезпечує достовірні результати аналізу змін станів дефектів на поверхнях елементів конструкцій.

Отримані у дисертаційній роботі наукові та практичні результати можуть бути використані під час проектування та дослідження кіберфізичних систем як для оптимізації загальної процедури досліджень зображень дефектів, так і для розробки методик підтримки прийняття рішень у відповідних комплексних інформаційно-аналітичних системах.

**СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ**

1. Шувар Р., Продивус А., Юзевич В., Огірко І., Огірко О., Ковтко Р., Мисюк Р. Інформаційні технології та загрози у кіберфізичних системах для відображення інформації у підземних металокопструкціях з дефектами. *Штучний інтелект*. 2021. Вип. 26 (1). С. 85-94.
2. R. Mysiuk, V. Yuzevych, I. Mysiuk. Data processing and visualization for corrosion detection. *Electronics and Information Technologies*. 2021. Vol. 16. P. 50–60.
3. R. Mysiuk, V. Yuzevych, I. Mysiuk, I. Ohirko. Detecting cracks in concrete based on images using Amazon Web Service Rekognition. *Electronics and Information Technologies*. 2022. Vol. 20. P. 18–29.
4. Mysiuk R., Yuzevych V., Mysiuk I. Api test automation of search functionality with artificial intelligence. *Artificial Intelligence*. 2022. Vol. 27, No 1. P. 269–274.
5. R. Mysiuk, V. Yuzevych. IoT-based solution for detection defects in infrastructure objects using Raspberry PI. *Electronics and Information Technologies*. 2023. No. 21. P. 45–56.
6. R. Mysiuk, I. Mysiuk, G. Pawlowski. Data Modeling of Physical-Mechanical Processes in Nanoconcrete with the Ensemble of Pores. *Path of Science*. 2022. V. 8, No. 9. P. 1001-1009.
7. R. Mysiuk, I. Mysiuk, V. Yuzevych, G. Pawlowski. Determining the Place of Depressurization of Underground Pipelines (Gas Pipelines): New Solutions in Industry based on Thermal Image Analysis Using Computer Vision. *Path of Science*. 2022. Vol. 8, No. 10. P. 1001-1010.
8. Mysiuk R., Yuzevych V. Recover Data about Detected Defects of Underground Metal Elements of Constructions in Amazon Elasticsearch Service. *Path of Science*. 2023. Vol. 9, No. 1. P. 1011-1019.
9. Mysiuk R. Pipeline Damage Assessment Based on Corrosion Segmentation Using JetRacer Kit. *Path of Science*. 2023. Vol. 9, No. 8. P. 6001-6006.

10. R.V. Mysiuk, V.M. Yuzevych, M.F. Yasinskyi, S.V. Kniaz, Z.A. Duriagina, V.V. Kulyk. Determination of conditions for loss of bearing capacity of underground ammonia pipelines based on the monitoring data and flexible search algorithms. *Archives of Materials Science and Engineering*. 2022. Vol. 115, No. 1. P. 13–20.
11. R. Mysiuk, V. Yuzevych, B. Koman and M. Yasinskyi. High Availability System for Monitoring Material Degradation Processes at the Concrete-polymer Interface. *2022 12<sup>th</sup> International Conference on Advanced Computer Information Technologies (ACIT)*, Spisska Kapitula, Slovakia, 26–28 September 2022. P. 415–418.
12. R. Dzhala, V. Yuzevych, R. Mysiuk, V. Brych, R. Skrynkovskyy, V. Lozovan, Y. Tyrkalo Simulation of Corrosion Fracture of Nano-Concrete at the Interface with Reinforcement Taking into Account Temperature Change. *CEUR Workshop Proceedings. MoMLeT+DS 2022: 4<sup>th</sup> International Workshop on Modern Machine Learning Technologies and Data Science*, November, 25–26, 2022, Leiden-Lviv, The Netherlands-Ukraine, Vol. 3312, P. 123–133.
13. R. Mysiuk, I. Mysiuk, G. Pawlowski, V. Yuzevych, M. Yasinskyi and Y. Tyrkalo Video-based Concrete Road Damage Assessment Using JetRacer Kit. *2023 17<sup>th</sup> International Conference on the Experience of Designing and Application of CAD Systems (CADSM)*, Jaroslaw, Poland, February 22–25, 2023. P. 1–4.
14. R. Mysiuk, V. Yuzevych, B. Koman, Y. Tyrkalo, O. Farat, I. Mysiuk, L. Harasym Detection of Structure Changes in Lightweight Concrete with Nanoparticles Using Computer Vision Methods in the Construction Industry. *Proceedings of Eighth International Congress on Information and Communication Technology (ICICT)*, Lecture Notes in Networks and Systems, vol 694. Springer, Singapore. United Kingdom, London, 20–23 February 2023 – P. 339–348.
15. R. Mysiuk, V. Yuzevych, I. Mysiuk. Y. Tyrkalo, A. Pavlenchyk, V. Dalyk. Detection of Surface Defects Inside Concrete Pipelines Using Trained Model on JetRacer Kit. *2023 IEEE 13th International Conference on Electronics and Information Technologies (ELIT)*, Lviv, Ukraine, September 26 – 28, 2023.

16. Mysiuk R. Information technology for risk-based pipeline resource assessment. *International scientific journal "Internauka"*. 2024. №1.
17. Mysiuk R. Towards Information Flows in Recognition and Prediction Tasks with Internet of Things. *Path of Science*. 2024. Vol. 10, No.1. P.1001-1004.
18. Мисюк Р., Юзевич В. Система пошуку та відбору інформації про дефекти типу тріщин у базах знань. *Міжнародна наукова інтернет-конференція "Інформаційне суспільство: технологічні, економічні та технічні аспекти становлення (випуск 53)"*. Збірник тез доповідей: випуск 53 (м. Тернопіль, 16 листопада 2020 р.). Частина 2. Тернопіль. 2020. С. 57–60.
19. R. Mysiuk, V. Yuzevych, B. Koman. Quality control of underground metal constructions with corrosion fatigue. *Quality management in education and industry: experience, problems and perspectives: proceedings of the V international scientific practical conference*, May 20–21, 2021. Lviv: «Piramida», 2021. P. 161– 163
20. Мисюк Р., Юзевич В. Гнучкий алгоритм у системі пошуку та відбору інформації про дефекти типу тріщин. *Міжнародна наукова інтернет-конференція "Інформаційне суспільство: технологічні, економічні та технічні аспекти становлення (випуск 55)"*. Збірник тез доповідей: випуск 55 (м. Тернопіль, 9 лютого 2021 р.). Тернопіль. 2021. С. 41–44.
21. R. Mysiuk, V. Yuzevych. Diagnosis of corrosion fatigue in underground metal constructions using genetic algorithms. *III International Scientific and Practical Internet Conference "Mathematics and Informatics in Higher Education: Challenges of Modernity"*, dedicated to the memory of Professors O. A. Pankov and V. S. Trokhymenko (Vinnytsia, May 20–21, 2021): book of abstracts. [Electronic network scientific publication], Vinnytsia, 2021, P. 93–96.
22. R. Mysiuk. Statistical models of pipe configurations for assessment of defects in infrastructure objects. *Proceedings of the XVI International Scientific and Practical Conference "Information Technologies and Automation – 2023"*. Odesa, October 19–20, 2023. Odesa, ONTU Publishing House, 2023, P. 69–71.

23. Features importance in statistical models for detecting material cracks. *Quality management in education and industry: experience, problems and perspectives: proceedings of the VI International Scientific and Practical Conference*, November 16 – 17, 2023, P. 152–153.
24. R. Mysiuk, I. Mysiuk, V. Yuzevych, R. Shuvar, S. Tsyuh and N. Pavlenchyk Predictive Analysis of Macro Defects in Engineering Structures Using Machine Learning Technologies. *International Conference on Optimization and Data Science in Industrial Engineering*, November 16-17, 2023, Istinye University, Istanbul, Turkey – P.111
25. 4 Emerging Technologies You Need to Know About. *Gartner*. URL: <https://www.gartner.com/en/articles/4-emerging-technologies-you-need-to-know-about>
26. State of the Art in Defect Detection Based on Machine Vision / Z. Ren et al. *International Journal of Precision Engineering and Manufacturing-Green Technology*. 2021
27. В. О. Гороховатський, І. С. Творошенко. Методи інтелектуального аналізу та оброблення даних : навч. посіб.. М-во освіти і науки України, Харків. нац. ун-т радіоелектроніки. *Харків : ХНУРЕ*, 2021. 92 с.
28. Lu, J., Lee, S.-H. Real-Time Defect Detection Model in Industrial Environment Based on Lightweight Deep Learning Network. *Electronics*. 2023. Vol. 12. P.4388.
29. Real-Time Defect Detection for Metal Components: A Fusion of Enhanced Canny–Devernay and YOLOv6 Algorithms / H. Wang та ін. *Applied Sciences*. 2023. Т. 13, № 12. С. 6898.
30. Zhao S., Kang F., Li J. Non-Contact Crack Visual Measurement System Combining Improved U-Net Algorithm and Canny Edge Detection Method with Laser Rangefinder and Camera. *Applied Sciences*. 2022. Т. 12, № 20. С. 10651.
31. Saberironaghi, A., Ren, J., El-Gindy, M. Defect Detection Methods for Industrial Products Using Deep Learning Techniques: A Review. *Algorithms*. 2023. Т.16. С. 95.

32. M. Dovbnych, M. Plechawska–Wójcik. A comparison of conventional and deep learning methods of image classification. *Journal of Computer Sciences Institute*. 2021. Вип.21. С. 303-308.
33. I. Konovalenko., P. Maruschak, J. Brezinová, O. Prentkovskis, Brezina, J. Research of U-Net-Based CNN Architectures for Metal Surface Defect Detection. *Machines*. 2022. Вип.10. С.327.
34. Natarajan, V., Hung, T.Y., Vaikundam, S., Chia, L.T. Convolutional networks for voting-based anomaly classification in metal surface inspection. *In Proceedings of the 2017 IEEE International Conference on Industrial Technology (ICIT)*, Toronto, ON, Canada, 22–25 March 2017. P. 986–991.
35. A deep-learning-based approach for fast and robust steel surface defects classification / G. Fu та ін. *Optics and Lasers in Engineering*. 2019. Т. 121. С. 397–405
- 36 Surface defect classification of steels with a new semi-supervised learning method / Di, H та ін. *Optics and Lasers in Engineering*. 2019. Вип.117. С. 40–48.
37. High Efficient Single-stage Steel Surface Defect Detection. / Akhyar, F та ін. *In Proceedings of the 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, Taipei, Taiwan, 18–21 September 2019. С. 1–4.
38. K. Li, X. Wang, L. Ji. Application of Multi-Scale Feature Fusion and Deep Learning in Detection of Steel Strip Surface Defect. *In Proceedings of the 2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM)*, Dublin, Ireland, 17–19 October 2019. С. 656–661.
39. Cha Y.-J., Choi W., Büyüköztürk O. Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks. *Computer-Aided Civil and Infrastructure Engineering*. 2017. Т. 32, № 5. С. 361–378..
40. Yu Z. YOLO V5s-based Deep Learning Approach for Concrete Cracks Detection. *SHS Web of Conferences*. 2022. Т. 144. С. 03015.

41. Welding defect classification based on convolution neural network (CNN) and Gaussian kernel / A. Khumaidi та ін. *In Proceedings of the 2017 International Seminar on Intelligent Technology and Its Applications (ISITIA)*. Surabaya, Indonesia, 28–29 August 2017. С. 261–265.
42. Detection, Localisation and Assessment of Defects in Pipes Using Guided Wave Techniques: A Review / Aidin Ghavamian та ін. *Sensors*. 2018. Т. 18, № 12. С. 4470.
43. Obshta A., Yuzevych V., Pohrebniak A., Mysiuk R., Chorniy B. Diagnostics of oil leaks caused by malicious damage to the linear part of oil pipelines: innovative solutions for the oil industry. *International scientific journal "Internauka"*. 2024. №2.
44. В. Гошовський, В. Дзіковський, Р. Мисюк, В. Рабик, І. Сасовець. Система збирання інформації на основі мікрокомп'ютера Raspberry PI. *Електроніка та інформаційні технології*. 2017. Вип. 8. С. 102–110.
45. A. Obshta, Y. Biliak, V. Shugai. Cyber-Physical System for Diagnostic Along the Controlled Section of the Oil Pipeline. *Advances in CyberPhysical Systems*. 2023 . Вип. 8 (1). С. 66–73.
46. Residual Life and Reliability Assessment of Underground RC sanitary Sewer Pipelines Under Uncertainty. *LSU Scholarly Repository*. URL: [https://repository.lsu.edu/transet\\_pubs/126](https://repository.lsu.edu/transet_pubs/126)
47. Hansen, P. Visual mapping for natural gas pipe inspection / Hansen, P [та ін.] // *The International Journal of Robotics Research* – 2015. – Vol. 34(4–5). – P. 532–558.
48. PyTorch. *PyTorch*. URL: <https://pytorch.org/>
49. ONNX Runtime | Inference. *ONNX Runtime | Home*. URL: <https://onnxruntime.ai/inference>
50. GitHub - kxnhha/crack\_segmentation: This repository contains code and dataset for the task crack segmentation using two architectures UNet\_VGG16, UNet\_Resnet and DenseNet-Tiramusu. *GitHub*. URL: [https://github.com/kxnhha/crack\\_segmentation](https://github.com/kxnhha/crack_segmentation)



51. vgg16 – Torchvision main documentation. *PyTorch*. URL: <https://pytorch.org/vision/main/models/generated/torchvision.models.vgg16.html>
52. Deepika H C. An Overview of You Only Look Once: Unified, Real-Time Object Detection. *International Journal for Research in Applied Science and Engineering Technology*. 2020. Т. 8, № 6. С. 607–609.
53. Roboflow: Give your software the power to see objects in images and video // *Roboflow: Give your software the power to see objects in images and video*. URL: <https://roboflow.com/>
54. Ultralytics. Home. *Home - Ultralytics YOLOv8 Docs*. URL: <https://docs.ultralytics.com/>
55. Б. М. Гавриш, О. В. Тимченко, М. П. Кляп. Алгоритми розпізнавання об'єктів на зображенні. *Поліграфія і видавнича справа*. 2022. Вип. 1 (83). С. 47-58.
56. Evaluation of the effectiveness of different image skeletonization methods in biometric security systems / Nazarkevych M., Dmytruk S., Hrytsyk V., Vozna O., Kuza A., Shevchuk O., Voznyi Y., Maslanych I., Sheketa V. *International Journal of Sensors, Wireless Communications and Control*. 2021. Вип. 11 (5). С. 542 - 552
57. A Review of Image Detection, Recognition and Classification with the Help of Machine Learning and Artificial Intelligence / Riya Kumari та ін. *SSRN Electronic Journal*. 2020. URL: <https://doi.org/10.2139/ssrn.3611339>
58. W. Nash, L. Zheng, N. Birbilis. Deep learning corrosion detection with confidence *npj Materials Degradation*. 2022. Т. 6, № 1.
59. Civil Infrastructure Damage and Corrosion Detection: An Application of Machine Learning / Hafiz Suliman Munawar та ін. *Buildings*. 2022. Т. 12, № 2. С. 156.
60. Неруйнівний контроль стану поверхонь, уражених корозійними пітингами / Б. П. Русин та ін. *Фізико-хімічна механіка матеріалів*. 2013. Т. 49, № 4. С. 90-96.
61. Image-Based Corrosion Detection in Ancillary Structures / A. Das та ін. *Infrastructures*. 2023. Т. 8, № 4. С. 66.

62. Detection and quantitative assessment of corrosion on pipelines through image analysis / V. Bondada та ін. *Procedia Computer Science*. 2018. Т. 133. С. 804–811
63. Pipeline Inspection for Corrosion using a Mobile Robotic Systems / Godwin Itseoritseagba та ін. *International Journal of Robotic Engineering*. 2015. Т. 1, № 1.
64. A review on various types of in-pipe inspection robot / Ankush Verma та ін. *Materials Today: Proceedings*. 2021.
65. Z. Gao, Y. Qiu, J. Zhou Paper Defect Detection Algorithm Based on the Mathematical Morphology and Computer Vision. *Journal of Physics: Conference Series*. 2022. Т. 2365, № 1. С. 012045.
66. S. Vani, H. Singh The Performance Analysis of Edge Detection Algorithms for Image Processing Based on Improved Canny Operator. *International Journal of Computer Trends & Technology*. 2020. Т. 68, № 10. С. 29–34..
67. Сусліков Л.М., Студеняк І.П. Неруйнівні методи контролю: *Навчальний посібник*. Ужгород: Видавництво УжНУ. 2016. 192 с. URL: <https://www.uzhnu.edu.ua/uk/infocentre/get/60790>
68. AI Thermometer. *Hackster.io*. URL: <https://www.hackster.io/tomasz-lewicki/ai-thermometer-2bacb4>
69. DeepEdgeBench: Benchmarking Deep Neural Networks on Edge Devices. / S. P. Baller та ін. *arXiv:2108.09457*. 2021. URL: <https://doi.org/10.48550/arXiv.2108.09457>
70. YOLOv7 with TensorRT on Jetson Nano (with Python script example). *Medium*. URL: <https://medium.com/@jurespeh/yolov7-with-tensorrt-on-jetson-nano-with-python-script-example-63099fa7c8a5>
71. GitHub - Qengineering/YoloV5-ncnn-Jetson-Nano: YoloV5 for Jetson Nano. GitHub. URL: <https://github.com/Qengineering/YoloV5-ncnn-Jetson-Nano>
72. YOLOv5 is Here!. *Towards Data Science*. URL: <https://towardsdatascience.com/yolo-v5-is-here-b668ce2a4908>

73. Jetson Nano Benchmarks - APIs - ximea support. *XIMEA - XIMEA*. URL: [https://www.ximea.com/support/wiki/apis/Jetson\\_Nano\\_Benchmarks](https://www.ximea.com/support/wiki/apis/Jetson_Nano_Benchmarks)
74. Google Colaboratory. *Google Colab*. URL: <https://colab.research.google.com/github/roboflow-ai/notebooks/blob/main/notebooks/train-yolov8-object-detection-on-custom-dataset.ipynb#scrollTo=Wjc1ctZykYuf>
75. Проект Плану відновлення України. Матеріали робочої групи “Аудиту збитків, понесених внаслідок війни”. Національна рада з відновлення України від наслідків війни. *Урядовий портал*. URL: <https://www.kmu.gov.ua/storage/app/sites/1/recoveryrada/ua/audit-of-war-damage.pdf>
76. Distribution, Transmission & Gathering, LNG, and Liquid Accident and Incident Data | PHMSA. *Pipeline and Hazardous Materials Safety Administration*. URL: <https://www.phmsa.dot.gov/data-and-statistics/pipeline/distribution-transmission-gathering-lng-and-liquid-accident-and-incident-data>
77. Distribution, Transmission & Gathering, LNG, and Liquid Accident and Incident Data | PHMSA. *Pipeline and Hazardous Materials Safety Administration*. URL: <https://www.phmsa.dot.gov/data-and-statistics/pipeline/distribution-transmission-gathering-lng-and-liquid-accident-and-incident-data>
78. Mechanical Fitting Failure Data from Gas Distribution Operators | PHMSA. *Pipeline and Hazardous Materials Safety Administration*. URL: <https://www.phmsa.dot.gov/data-and-statistics/pipeline/mechanical-fitting-failure-data-gas-distribution-operators>
79. 2023 - Download NBI ASCII files - National Bridge Inventory - Bridge Inspection - Safety Inspection - Bridges & Structures - Federal Highway Administration. *Federal Highway Administration*. URL: <https://www.fhwa.dot.gov/bridge/nbi/ascii2023.cfm>

80. Lee, G. C., Mohan, S. B., Huang, C., Fard, B. N. Study of U.S. Bridge Failures (1980-2012). *National Technical Reports Library*. URL: <https://ntrl.ntis.gov/NTRL/dashboard/searchResults/titleDetail/PB2016101915.xhtml>
81. Хоменко А., М. Гомеля, О. Степова. Вплив показників ґрунтового середовища на корозійний стан нафтопроводів. *Матеріали міжнародної науково-практичної конференції "Екологія. Людина. Суспільство"*. 2021. С. 263–267.
82. Climate data for cities worldwide. *Climate data for cities worldwide*. URL: <https://en.climate-data.org/>
83. Surface Crack Detection Using Data Mining and Feature Engineering Techniques. / A. Chordia та ін. *2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON)*. Kuala Lumpur, Malaysia . 2021. P. 1–7.
84. Comparative Analysis of Machine Learning Models for Predicting Crack Propagation under Coupled Load and Temperature / I. Omar та ін. *Applied Sciences*. 2023. Т. 13, № 12. С. 7212.
85. A review on pipeline failures, defects in pipelines and their assessment and fatigue life prediction methods / S. Vishnuvardhan та ін. *International Journal of Pressure Vessels and Piping*. 2022. С. 104853.
86. Degradation Assessment and Failure Prevention of Pipeline Systems. *Lecture Notes in Civil Engineering* In: Bolzon G., Gabetta G., Nykyforchyn H. (eds). Springer, Cham. 2021. Т. 102. С. 203216.
87. Failure Risk Prediction Model for Girth Welds in High-Strength Steel Pipeline Based on Historical Data and Artificial Neural Network / Ke Wang та ін. *Processes*. 2023. Т. 11, № 8. С. 2273
88. Improving pipe failure predictions: Factors affecting pipe failure in drinking water networks / Neal Andrew Barton та ін. *Water Research*. 2019. Т. 164. С. 114926.
89. Ванькович Л. Я., Мисюк Р. В., Добуш Т. І., Бобко О. В., Коник О. В., Шиян С. М., Цвик С. Р. Вплив інформаційних технологій на розвиток підприємства.

*Міжнародний науковий журнал "Інтернаука". Серія: "Економічні науки". 2023. №9.*

90. Юзевич В. М., Павловські Г., Павленчик А. О., Мисюк Р. В., Тиркало Ю. Є., Ільчишин М. З. Оптимізація управлінського рішення щодо асортиментної політики підприємства із використанням математичного моделювання в умовах ризику. *Міжнародний науковий журнал "Інтернаука". Серія: "Економічні науки". 2023. №2.*

91. Olenych I., Olenych Y., Kostruba A., Pryima Y. Fractal Analysis of Porous Structures Using a Fuzzy Logic System. *2019 11th International Scientific and Practical Conference on Electronics and Information Technologies (ELIT 2019)*. 2019. С. 97-101

92. Oleksii Dudnyk, Zoia Sokolovska. Forecasting development trends in the information technology industry using fuzzy logic. *Eastern-European Journal of Enterprise Technologies*. 2023. Т. 1, № 13 (121). С. 74–85.

93. H. A. Al-Jamimi. Toward comprehensible software defect prediction models using fuzzy logic. *2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS)*. Beijing, China. 2016. P. 127–130.

94. R. Jain, R. S. Sharma, Predicting Severity of Cracks in Concrete using Fuzzy Logic *2018 International Conference on Recent Innovations in Electrical, Electronics & Communication Engineering (ICRIEECE)* . Bhubaneswar, India. 2018. С. 2976–2976.

95. M. Victoria Biezma, Diego Agudo, Gonzalo Barron, A Fuzzy Logic method: Predicting pipeline external corrosion rate, *International Journal of Pressure Vessels and Piping*, 2018, Вып. 163, С 55–62, ISSN 0308-0161.

96. Pavlyshenko B. Bayesian regression approach for building and stacking predictive models in time series analytics. *Communications in Computer and Information Science*, 2020, Вып. 1158, С. 486 - 500.

97. Москаленко, В. В., Кріпак, С. А. Дослідження методів машинного навчання для аналізу та прогнозування закупівельних даних. *Таврійський науковий вісник. Серія: Технічні науки*, 2023, Вип. 4, С. 61-68.
98. Boyko, Nataliya. Research into machine learning algorithms for the construction of mathematical models of multimodal data classification problems. *Computational Problems of Electrical Engineering*. 2021. Вип.11. С. 1-11.
99. Застосування методів машинного навчання для статистичного аналізу та прогнозування кіберспортивної галузі / О. Кузьменко та ін. *Measuring and computing devices in technological processes*, 2021. № 1, С. 126–132.
100. Jason E Black, Jacqueline K Kueper, Tyler S Williamson, An introduction to machine learning for classification and prediction / Jason E Black та ін. *Family Practice*, 2023, Вип. 40, № 1, С. 200–204.
101. Feature Importance and Visualization of Tree Models. *Medium*. URL: <https://medium.com/chinmaygaikwad/feature-importance-and-visualization-of-tree-models-d491e8198b0a>
102. How to Calculate Feature Importance With Python. *Machine Learning Mastery*. URL: <https://machinelearningmastery.com/calculate-feature-importance-with-python/>
103. Machine Learning Metrics in simple terms. *Medium*. URL: <https://medium.com/analytics-vidhya/machine-learning-metrics-in-simple-terms-d58a9c85f9f6>
104. Labeling Data with Pandas. *Towards Data Science*. URL: <https://towardsdatascience.com/labeling-data-with-pandas-9e573ce59c42>
105. Anton N., Nedelcu A. Analysis of information and information flow in technological processes. Method of transmitting information unaltered. *IOP Conference Series: Materials Science and Engineering*. 2019. Т. 564. С. 012094.
106. The analysis of information flow interdependencies within projects / H. Bashir та ін. *Production Planning & Control*. 2020. С. 1–17.

107. Anthony B. Information Flow Analysis of a Knowledge Mapping-Based System for University Alumni Collaboration: a Practical Approach. *Journal of the Knowledge Economy*. 2020. C. 756–787.
108. Application of Internet of Things in Image Processing / N. Kumar та ін. 2022 *IEEE Delhi Section Conference (DELCON)*, New Delhi, India, 2022, C. 1-5.
109. A Review: Image Processing Techniques' Roles towards Energy-Efficient and Secure IoT / A. M. Al-Ghaili та ін. *Applied Sciences*. 2023. T. 13, № 4. C. 2098.
110. Petroleum pipeline monitoring using an internet of things (IoT) platform / E. N. Aba та ін. *SN Applied Sciences*. 2021. T. 3, № 2.
111. Water Pipeline Leakage Monitoring System based on Internet of Things / N. N. Che та ін. *Journal of Physics: Conference Series*. 2021. T. 1962, № 1. C. 012025.
112. Kashyap M., Sharma D. V. Design and Performance Evaluation of Mqtt Based Stimulus Communication System for Iot. *SSRN Electronic Journal*. 2022.
113. Muñoz Castañeda Á. L., Mata J. A. A., Aláiz-Moretón H. Characterization of threats in IoT from an MQTT protocol-oriented dataset. *Complex & Intelligent Systems*. 2023.
114. Security Analysis of the MQTT-SN Protocol for the Internet of Things / J. Roldán-Gómez та ін. *Applied Sciences*. 2022. T. 12, № 21. C. 10991.
115. Abbas N., Shujaa M. I., Zghair E. M. IoT Based Image Processing Filters. *Journal La Multiapp*. 2021. T. 2, № 4. C. 1–13.
116. Deep Learning-Based Smart Colored Fabric Defect Detection System / S. Barua та ін. *Applied Computer Vision and Image Processing*. Singapore, 2020. C. 212–219.
117. Toward Efficient Image Recognition in Sensor-Based IoT: A Weight Initialization Optimizing Method for CNN Based on RGB Influence Proportion / Z. Deng та ін. *Sensors*. 2020. T. 20, № 10. C. 2866.
118. Bulk API | Elasticsearch Guide [8.12] | Elastic. *Elasticsearch Platform – Find real-time answers at scale* | Elastic. URL: <https://www.elastic.co/guide/en/elasticsearch/reference/current/docs-bulk.html>

119. NVIDIA AI IOT. *GitHub: Let's build from here · GitHub*. URL: [https://github.com/NVIDIA-AI-IOT/jetracer/blob/master/notebooks/basic\\_motion.ipynb](https://github.com/NVIDIA-AI-IOT/jetracer/blob/master/notebooks/basic_motion.ipynb)
120. How to Train a Machine Learning Model as a REST API and Build a Spam Classifier (Part 1). *Medium*. URL: <https://medium.com/building-the-open-data-stack/how-to-train-a-machine-learning-model-as-a-rest-api-and-build-a-spam-classifier-part-1-385f3b554b18>
- 121 Nguyen T. T. Machine learning-based model for predicting concrete compressive strength. *International Journal of GEOMATE*. 2021. T. 20, № 77. URL: <https://doi.org/10.21660/2020.77.j2019>.
122. Serve a machine learning model using Sklearn, FastAPI, and Docker. *Medium*. URL: <https://medium.com/analytics-vidhya/serve-a-machine-learning-model-using-sklearn-fastapi-and-docker-85aabf96729b>
123. Wrapping Machine Learning Model in an API — Using FASTAPI. *Medium*. URL: <https://medium.com/@sylvesterjnr87/wrapping-machine-learning-model-in-an-api-using-fastapi-aeb07b131246>
124. django-nvd3. *PyPI*. URL: <https://pypi.org/project/django-nvd3/>
125. Connecting | Elasticsearch Python Client [8.12] | Elastic. *Elasticsearch Platform – Find real-time answers at scale | Elastic*. URL: <https://www.elastic.co/guide/en/elasticsearch/client/python-api/current/connecting.html>
126. Python in Beats | Beats Developer Guide [7.17] | Elastic. *Elasticsearch Platform – Find real-time answers at scale | Elastic*. URL: <https://www.elastic.co/guide/en/beats/devguide/7.17/python-beats.html>
127. python-elastic-logstash. *PyPI*. URL: <https://pypi.org/project/python-elastic-logstash/>



128. Salgado A. Bringing Elastic to Robotics Developers. *Elastic Blog*. URL: <https://www.elastic.co/blog/bringing-elastic-to-robotics-developers-integrating-with-ros>
129. Жерновий Ю. В. Імітаційне моделювання систем масового обслуговування: Практикум. Львів: Видавничий центр ЛНУ імені Івана Франка, 2007. 307 с.
130. Uryvsky L., Moshynska A., Osypchuk S. Analysis of digital signal quality indicators based on analytic and stochastic modeling. *2017 International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo)*, м. Odessa, 11–15 верес. 2017 р. 2017. URL: <https://doi.org/10.1109/ukrmico.2017.8095414>
131. Кучук Н. Г. Метод розрахунку максимальних інтенсивностей інформаційних потоків у гіперконвергентній системі. *Системи управління, навігації та зв'язку*. 2019. Вип. 4. С. 53-56. URL: [http://nbuv.gov.ua/UJRN/suntz\\_2019\\_4\\_13](http://nbuv.gov.ua/UJRN/suntz_2019_4_13).
132. Index stats API | Elasticsearch Guide [8.12] | Elastic. *Elasticsearch Platform – Find real-time answers at scale | Elastic*. URL: <https://www.elastic.co/guide/en/elasticsearch/reference/8.12/indices-stats.htm>
133. Корнієнко І.В., Корнієнко С.П. Шевага В.В., Руденко О.В., Жирна О.В. Функціональне моделювання й аналіз інформаційних потоків системи випробувань озброєння та військової техніки. *Збірник наукових праць Державного науково-дослідного інституту випробувань і сертифікації озброєння та військової техніки*. 2021. Вип. № 4(10). URL: <https://doi.org/10.37701/dndivsovt.10.2021.09>
134. Гущак Ж. М., Добош У. П. Дефекти магістральних газопроводів. *Міжвузівський збірник "НАУКОВІ НОТАТКИ"*. Луцьк, 2013. Випуск № 40. С. 339–342.
135. Lozovan V., Skrynkovsky R., Yuzevych V., Yasinskyi M., Pawlowski, G. Forming the toolset for development of a system to control quality of operation of underground pipelines by oil and gas enterprises with the use of neural networks.

*Eastern-European Journal of Enterprise Technologies*. 2019. Vol. 2, No 5 (98). P. 41–48. URL: <http://dx.doi.org/10.15587/1729-4061.2019.161484>

136. Ширмовська Н. Г., Николайчук Я. М. Моделювання процесів діагностування станів складних об'єктів управління комп'ютерних систем // *Вісник Національного університету "Львівська політехніка". Комп'ютерні системи та мережі*. 2010. № 688. С. 219-224.

137. Лавріщева К. М., Коваль Г. І., Коротун Т. М. Підходи до інженерії якості сімейств програмних систем. *Проблеми програмування*. 2008. № 2–3 (спецвипуск УкрПРОГ'2008). С. 219–228.

138. Русин Б. П., Іванюк В.Г., Капшій О.В., Ануфрієва Н. П., Похмурський А. Ю. Оцінка глибини пітінга за зображеннями поверхні матеріалу. *Радіoeлектроніка та інформатика*, 2010, № 1. С. 83-91.

139. Сиротюк А. М., Капшій О. В., Дмитрах І. М., Русин Б. П., Білий О. Л. Експертна система оцінювання ризику руйнування конструктивних елементів трубопроводів теплоенергетичної системи. *Методи та прилади контролю якості*, 2013, № 2 (31). С. 126-135.

140. Поляков С. Г., Клименко А. В., Коваленко С. Ю. Система корозійного моніторингу трубопроводів. *Наука та інновації*. 2010. Т. 6, № 5. С. 25-28.

141. Kaeshe H. Die Korrosion der Metalle. Physikalisch-chemische Prinzipien und aktuelle Probleme. *Berlin-Heidenberg-New York: Springer-Verlag*, 1979. 400 p.

142. Валяшек В. Б., Каплун А. В., Юзевич В. М. Математичне та комп'ютерне моделювання фізичних характеристик матеріалу у вершині тріщини з урахуванням ефекту зміцнення. *Комп'ютерно-інтегровані технології: освіта, наука, виробництво*. 2015. № 18. С. 97-104.

143. Юзевич В. М., Огірко О. І. Інформаційна технологія оцінювання станів об'єктів із сталі в морському середовищі із сірководнем з використанням алгебри алгоритмів. *Наукові записки (Українська академія друкарства)*. 2012. № 4 (41). С. 160–172.

144. Юзевич В. М., Байцар Р. І., Гунькало А. В. СУЯ: моделювання управління процесами. *Вісн. Нац. ун-ту "Львів. політехніка"*. 2007. № 574. С. 123-126.
145. Копитко М. І. Комплексне забезпечення економічної безпеки підприємств (на матеріалах підприємств транспортного машинобудування): *дис. на здобуття наук. ступеня докт. економіч. наук: спец. 21.04.02 "Економічна безпека суб'єктів господарської діяльності"*. Київ, 2015. 478 с.
146. Qualimetric Analysis of Pipelines with Corrosion Surfaces in the Monitoring System of Oil and Gas Enterprises / V. Yuzevych, N. Pavlenchuk, O. Zaiats, N. Heorhiadi, V. Lakiza. *International Journal of Recent Technology and Engineering (IJRTE)*, 2020. Vol. 9, No. 1. P. 1145–1150.
147. Yuzevych, V. M., Lozovan, V. P. Influence of Mechanical Stresses on the Propagation of Corrosion Cracks in Pipeline Walls. *Materials Science*, 2022. Vol. 57, No. 4. P. 539-548.
148. Predictive maintenance: Weibull analysis with Python. *Medium*. URL: <https://m-berta.medium.com/predictive-maintenance-weibull-fit-with-python-09747707cbca>
149. How to Analyze the Diagnostic Performance of a New Test? Explained with Illustrations / D. Dhamnetiya та ін. *Journal of Laboratory Physicians*. 2021.

**Додаток А**  
**СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА**

**Наукові праці, в яких опубліковані основні наукові результати дисертації**

*Публікації у фахових виданнях України категорії Б:*

1. Шувар Р., Продивус А., Юзевич В., Огірко І., Огірко О., Ковтко Р., Мисюк Р. Інформаційні технології та загрози у кіберфізичних системах для відображення інформації у підземних металоконструкціях з дефектами // Штучний інтелект. – 2021. – Вип. 26 (1). – С. 85–94.
2. Mysiuk R., Yuzevych V., Mysiuk I. Data processing and visualization for corrosion detection // Electronics and Information Technologies. – 2021. – Vol. 16. – P. 50–60.
3. Mysiuk R., Yuzevych V., Mysiuk I., Ohirko I. Detecting cracks in concrete based on images using Amazon Web Service Rekognition // Electronics and Information Technologies. – 2022. – Vol. 20. – P. 18–29.
4. Mysiuk R., Yuzevych V., Mysiuk I. Api test automation of search functionality with artificial intelligence // Artificial Intelligence. – 2022. – Vol. 27, No 1. – P. 269–274.
5. Mysiuk R., Yuzevych V. IoT-based solution for detection defects in infrastructure objects using Raspberry PI // Electronics and Information Technologies. – 2023. – No. 21. – P. 45–56.

*Публікації у виданнях, які проіндексовані у міжнародній наукометричній базі Scopus:*

1. R.V. Mysiuk, V.M. Yuzevych, M.F. Yasinskyi, S.V. Kniaz, Z.A. Duriagina, V.V. Kulyk. Determination of conditions for loss of bearing capacity of underground ammonia pipelines based on the monitoring data and flexible search algorithms // Archives of Materials Science and Engineering – 2022. – Vol. 115, No. 1. – P. 13–20.
2. High Availability System for Monitoring Material Degradation Processes at the Concrete-polymer Interface / R. Mysiuk, V. Yuzevych, B. Koman and M. Yasinskyi //

2022 12<sup>th</sup> International Conference on Advanced Computer Information Technologies (ACIT) – Spisska Kapitula, Slovakia, 26–28 September 2022. – P. 415–418.

3. Simulation of Corrosion Fracture of Nano-Concrete at the Interface with Reinforcement Taking into Account Temperature Change / R. Dzhalá, V. Yuzevych, R. Mysiuk, V. Brych, R. Skrynkovskyy, V. Lozovan, Y. Tyrkalo // CEUR Workshop Proceedings. MoMLLeT+DS 2022: 4<sup>th</sup> International Workshop on Modern Machine Learning Technologies and Data Science, November, 25–26, 2022, Leiden-Lviv, The Netherlands-Ukraine – Vol. 3312 – P. 123–133.

4. Video-based Concrete Road Damage Assessment Using JetRacer Kit / R. Mysiuk, I. Mysiuk, G. Pawlowski, V. Yuzevych, M. Yasynskyi and Y. Tyrkalo // 2023 17<sup>th</sup> International Conference on the Experience of Designing and Application of CAD Systems (CADSM), Jaroslaw, Poland, February 22–25, 2023. – P. 1–4.

5. Detection of Structure Changes in Lightweight Concrete with Nanoparticles Using Computer Vision Methods in the Construction Industry / R. Mysiuk, V. Yuzevych, B. Koman, Y. Tyrkalo, O. Farat, I. Mysiuk, L. Harasym // Proceedings of Eighth International Congress on Information and Communication Technology (ICICT), Lecture Notes in Networks and Systems, vol 694. Springer, Singapore. United Kingdom, London, 20–23 February 2023 – P. 339–348.

6. Detection of Surface Defects Inside Concrete Pipelines Using Trained Model on JetRacer Kit / Mysiuk R., Yuzevych V., Mysiuk I. Tyrkalo Y., Pavlenchuk A., Dalyk V. // 2023 IEEE 13<sup>th</sup> International Conference on Electronics and Information Technologies (ELIT), Lviv, Ukraine, September 26 – 28, 2023. – P. 21–24.

***Публікації у наукових періодичних виданнях інших держав:***

1. Mysiuk R., Mysiuk I., Pawlowski G. Data Modeling of Physical-Mechanical Processes in Nanoconcrete with the Ensemble of Pores // Path of Science. – 2022. – V. 8, No. 9 – P. 1001–1009.

2. Mysiuk R., Mysiuk I., Yuzevych V., Pawlowski G. Determining the Place of Depressurization of Underground Pipelines (Gas Pipelines): New Solutions in Industry based on Thermal Image Analysis Using Computer Vision // Path of Science. – 2022. – Vol. 8, No. 10 – P. 1001–1010.
3. Mysiuk R., Yuzevych V. Recover Data about Detected Defects of Underground Metal Elements of Constructions in Amazon Elasticsearch Service // Path of Science. – 2023. – Vol. 9, No. 1 – P. 1011–1019.
4. Mysiuk R. Pipeline Damage Assessment Based on Corrosion Segmentation Using JetRacer Kit // Path of Science. – 2023. – Vol. 9, No. 8 – P. 6001–6006.
5. Mysiuk R. Information technology for risk-based pipeline resource assessment // International scientific journal "Internauka". – 2024. – №1.
6. Obshta A., Yuzevych V., Pohrebniak A., Mysiuk R., Chorniy B. Diagnostics of oil leaks caused by malicious damage to the linear part of oil pipelines: innovative solutions for the oil industry // International scientific journal "Internauka". — 2024. — №2.
7. Mysiuk R. Towards Information Flows in Recognition and Prediction Tasks with Internet of Things // Path of Science. – 2024. – Vol. 10, No.1 – P.1001–1004.

**Наукові праці, які засвідчують апробацію матеріалів дисертації:**

1. Мисюк Р., Юзевич В. Система пошуку та відбору інформації про дефекти типу тріщин у базах знань // Міжнародна наукова інтернет-конференція “Інформаційне суспільство: технологічні, економічні та технічні аспекти становлення (випуск 53)” / Збірник тез доповідей: випуск 53 (м. Тернопіль, 16 листопада 2020 р.). – Частина 2. – Тернопіль. – 2020. – С. 57–60.
2. Mysiuk R., Yuzevych V., Koman B. Quality control of underground metal constructions with corrosion fatigue // Quality management in education and industry:

experience, problems and perspectives: proceedings of the V international scientific practical conference, May 20–21, 2021 / Lviv: «Piramida», 2021. P. 161–163.

3. Мисюк Р., Юзевич В. Гнучкий алгоритм у системі пошуку та відбору інформації про дефекти типу тріщин // Міжнародна наукова інтернет-конференція “Інформаційне суспільство: технологічні, економічні та технічні аспекти становлення (випуск 55)” / Збірник тез доповідей: випуск 55 (м. Тернопіль, 9 лютого 2021 р.). – Тернопіль. – 2021. – С. 41–44.

4. Mysiuk R., Yuzevych V. Diagnosis of corrosion fatigue in underground metal constructions using genetic algorithms // III International Scientific and Practical Internet Conference “Mathematics and Informatics in Higher Education: Challenges of Modernity”, dedicated to the memory of Professors O. A. Pankov and V. S. Trokhymenko (Vinnytsia, May 20–21, 2021): book of abstracts. [Electronic network scientific publication], Vinnytsia, 2021, P. 93–96.

5. Mysiuk R. Statistical models of pipe configurations for assessment of defects in infrastructure objects // Information Technologies and Automation - 2023 / Proceedings of the XVI International Scientific and Practical Conference. Odesa, October 19–20, 2023. - Odesa, ONTU Publishing House – 2023 – P. 69–71.

6. Mysiuk R., Yuzevych V. Features importance in statistical models for detecting material cracks // Quality management in education and industry: experience, problems and perspectives: proceedings of the VI International Scientific and Practical Conference, November 16 – 17, 2023, - Lviv, Lviv Polytechnic Publishing House – 2023 – P. 152–153.

7. Mysiuk R., Mysiuk I., Yuzevych V., Shuvar R., Tsyuh S., Pavlenchuk N. Predictive Analysis of Macro Defects in Engineering Structures Using Machine Learning Technologies // International Conference on Optimization and Data Science in Industrial Engineering, November 16-17, 2023, Istinye University, Istanbul, Turkey – P.111.

***Відомості про апробацію результатів дисертації:***

1. Міжнародна наукова інтернет-конференція “Інформаційне суспільство: технологічні, економічні та технічні аспекти становлення (випуск 53)”, м. Тернопіль, 16 листопада 2020 року, форма участі – заочна.
2. XX Міжнародна науково-технічна конференція “Штучний інтелект та інтелектуальні системи” (AIIS’2020), м. Київ, 27 листопада 2020 року, форма участі – дистанційна, усна доповідь.
3. Міжнародна наукова інтернет-конференція “Інформаційне суспільство: технологічні, економічні та технічні аспекти становлення (випуск 55)”, м. Тернопіль, 9 лютого 2021 року, форма участі – заочна.
4. V Міжнародна науково-практична конференція "Управління якістю в освіті та промисловості: досвід, проблеми та перспективи" (QM-2021), м. Львів, 20–21 травня 2021 року, форма участі – дистанційна, усна доповідь.
5. III International Scientific and Practical Internet Conference “Mathematics and Informatics in Higher Education: Challenges of Modernity”, dedicated to the memory of Professors O. A. Pankov and V. S. Trokhymenko, Vinnytsia, Ukraine, May 20-21, 2021, форма участі – дистанційна, усна доповідь.
6. XXI International Scientific and Technical Conference “Artificial Intelligence and Intelligent systems” (AIIS’2021), м. Київ, 26–27 листопада 2021 року, форма участі – дистанційна, усна доповідь.
7. 2022 12<sup>th</sup> International Conference on Advanced Computer Information Technologies (ACIT) – Spisska Kapitula, SLOVAKIA, September 26–28, 2022, форма участі – дистанційна, усна доповідь.
8. 4<sup>th</sup> International Workshop on Modern Machine Learning Technologies and Data Science (MoMLLeT+DS 2022), November 25–26, 2022, Leiden-Lviv, The Netherlands-Ukraine, форма участі – дистанційна, усна доповідь.



9. 2023 17<sup>th</sup> International Conference on the Experience of Designing and Application of CAD Systems (CADSM), Jaroslaw, Poland, February 22–25, 2023, форма участі – дистанційна, усна доповідь.
10. 8<sup>th</sup> International Congress on Information and Communication Technology (ICICT 2023), London, United Kingdom, February 20–23, 2023, форма участі – дистанційна, усна доповідь.
11. 2023 IEEE 13th International Conference on Electronics and Information Technologies (ELIT), September 26–28, 2023, Lviv, Ukraine, форма участі – дистанційна, усна доповідь.
12. XVI International Scientific and Practical Conference “Information Technologies and Automation - 2023” (ITIA2023), Odesa, Ukraine, October 19–20, 2023, форма участі – дистанційна, усна доповідь.
13. VI International Scientific and Practical Conference “Quality management in education and industry: experience, problems and perspectives” (QM-2023), November 16–17, 2023, Lviv, Ukraine, форма участі – дистанційна, усна доповідь.
14. International Conference on Optimization and Data Science in Industrial Engineering (ODSIE2023), November 16–17, 2023, Istinye University, Istanbul, Turkey, форма участі – дистанційна, усна доповідь.
15. Наукові семінари кафедри системного проектування та факультету електроніки та комп’ютерних технологій Львівського національного університету імені Івана Франка у 2023 році, форми участі – очна, усні доповіді.
16. Звітні наукові конференції Львівського національного університету імені Івана Франка за 2021-2024 роки, форми участі – очна/дистанційна, усні доповіді.