

СУМСЬКИЙ ДЕРЖАВНИЙ УНІВЕРСИТЕТ
МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Кваліфікаційна наукова
праця на правах рукопису

ЗАРЕЦЬКИЙ МИКОЛА ОЛЕКСАНДРОВИЧ

УДК: 004.89.9:621.184.26(043.5)

ДИСЕРТАЦІЯ
МОДЕЛІ ТА МЕТОДИ ІНТЕЛЕКТУАЛЬНОЇ ІНФОРМАЦІЙНОЇ
ТЕХНОЛОГІЇ ОЦІНЮВАННЯ ФУНКЦІОНАЛЬНОГО СТАНУ ТРУБ
ВОДОВІДВЕДЕННЯ

Спеціальність – 122 Комп’ютерні науки
Галузь знань – 12 Інформаційні технології

Подається на здобуття наукового ступеня доктора філософії

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело

_____ Зарецький М.О.

Науковий керівник :
Москаленко В’ячеслав Васильович,
кандидат технічних наук, доцент

Суми – 2023

АНОТАЦІЯ

Зарецький М.О. Моделі та методи інтелектуальної інформаційної технології оцінювання функціонального стану труб водовідведення. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії з галузі знань 12 Інформаційні технології за спеціальністю 122 – Комп'ютерні науки. – Сумський державний університет, Суми, 2023.

Дисертація присвячена розв'язанню актуальної науково-прикладної задачі автоматизації процесу оцінювання функціонального стану трубопроводу водовідведення за даними відеоінспекції через розроблення інформаційної технології інтелектуального аналізу даних. Застосування розробленої інформаційної технології оцінювання функціонального стану труб водовідведення дозволяє підвищити точність автоматичного формування звітів про інспекцію.

Аналіз наукових джерел показав, що візуальні спостереження в інспекції труб водовідведення досить сильно відрізняються від зображень відкритих наборів даних, що знижує ефективність використання техніки переносу знань (transfer learning) для початкової ініціалізації моделі аналізу даних. При цьому неоднозначність кодування дефектів складної конфігурації, висока варіативність візуальних спостережень, рутинність і дороговартісність розмітки даних, обумовлюють обмеженість обсягу і високу незбалансованість розмічених навчальних даних, що обмежує ефективність традиційних методів навчання з учителем. Залежність кодування функціонального стану труби від зміни рівня води, орієнтації камери і дефектів під час відеоінспекції, наявність значної кількості артефактів та шумів обумовлюють необхідність врахування просторово-часової контекстної інформації. Тому одним із складних завдань, на дослідження якого спрямована дисертаційна робота, є створення інтелектуальної інформаційної технології машинного навчання для автоматичного аналізу даних відеоінспекції труб водовідведення за умов обмеженого обсягу навчальних даних та ресурсів, доступних для навчання та обслуговування системи. Водночас вирішення цього

завдання полягає в необхідності аналізу просторово-часової контекстної інформації.

Об'єктом дослідження є процес оцінювання функціонального стану труб водовідведення.

Предметом дослідження є моделі й методи інформаційної технології оцінювання функціонального стану труб водовідведення.

У дисертаційній роботі визначено актуальність застосування інформаційних технологій у галузі інспекції труб водовідведення. Для інспекції труб водовідведення використовують відеокамери, в тому числі тепловізійні, сонари, георадари, лазерні сканери, електросканери та їх комбінації з метою виявлення дефектів та оцінювання цілісності стінок труб та стану ґрунту навколо труб. На основі проведеного аналізу методів та підходів до інспекції труб водовідведення встановлено, що основним джерелом інформації є відеокамери, оскільки відеоінспекція вимагає найменше накладних витрат, а інформативність отриманих даних забезпечує виявлення більшості важливих типів дефектів. основна тенденція розвитку сучасних технологій аналізу даних відеоінспекції пов'язана з впровадженням технологій машинного зору та штучного інтелекту, хоча ці технології досі не надають достатнього рівня автоматизації через недостатню достовірність сформованих звітів. Таким чином, їх удосконалення є важливим напрямком досліджень. За результатами аналізу процесу відеоінспекції труб водовідведення обґрунтовано потребу у створенні нової інтелектуальної інформаційної технології оцінювання функціонального стану труб водовідведення шляхом використання ідей і методів композиції простих моделей, контрастного навчання, завадозахищеного кодування та інформаційно-екстремальної інтелектуальної технології (ІЕІ-технології). Таким чином, основний напрямок підвищення ефективності моделей оцінювання функціонального стану труб водовідведення полягає в адаптації вхідного математичного опису та побудові в процесі навчання класифікаційних вирішувальних правил, що призначені для розпізнавання контексту, розпізнавання дефектів в окремих контекстах та оцінювання зміни рівня води, шляхом максимізації критерію ефективності.

Уперше розроблено метод навчання ієрархічного екстрактора ознак, що на відміну від відомих, забезпечує підвищення функціональної ефективності вирішувальних правил за умов зашумленості та незбалансованості навчальних даних шляхом поєднання принципів переносу знань, сіамських мереж, контрастного самонавчання та самокоректуючих двійкових кодів. Запропонований метод машинного навчання забезпечує точність 98% для класифікаційного аналізу дефектів у контексті спостереження з орієнтацією камери вперед вздовж труби, що перевищує на 11% результат традиційного підходу до навчання з кінця-в-кінець з повнозв'язним шаром класифікації. Оскільки в контексті спостереження, що відповідає орієнтації камери на бічну стінку, присутня помітна незбалансованість даних, то ефективність класифікації дефектів оцінюється F1-метрикою і становить $F1=0,977$, що на 5% більше ніж в рамках традиційного підходу. Отримана точність класифікації дефектів в рамках запропонованого підходу перевищує на 6% результати отримані іншими авторами, і дозволяє отримати більшу точність звітів про функціональний стан труб водовідведення.

Удосконалено моделі екстракції ознакового опису спостережень шляхом поєднання локальної і просторово-часової візуальної контекстної інформації, що дозволяє підвищити інформативність ознакового опису відеоспостережень інспекції трубопроводу водовідведення за умов неповної визначеності. Встановлено перевагу використання 1D згорткової моделі порівняно з LSTM (Long Short-Term Memory) та GRU (Gated Recurrent Units) мережами для аналізу міжкадрових залежностей. Також емпірично доведено, що регулярна згорткова архітектура темпєральної мережі є найбільш ефективною під час аналізу міжкадрових залежностей відеоінспекції труб. Мікроусереднене значення F1-міри навченого в рамках запропонованого підходу екстрактора ознак для розпізнавання контексту спостережень з алфавіту 11 класів становить 0,92, що перевищує на 5% результат для моделі без врахування міжкадрових залежностей. Для алфавіту класів, що описує рівні води від 0% до 55% з кроком в 5% мікроусереднене значення F1-міри в рамках запропонованого підходу становить 0,879 і перевищує на 9% значення для моделі без врахування міжкадрових залежностей.

Набули подальшого розвитку моделі і алгоритми синтезу вирішувальних правил для класифікаційного аналізу спостережень, які основані на композиції моделей та оптимізації параметрів класифікаційних вирішувальних правил та гіперпараметрів екстрактора ознак за інформаційним критерієм. Тобто моделі з інформаційно-екстремальними вирішувальними правилами для розпізнавання контексту, рівня води та дефектів komponуються для оцінювання функціонального стану труби. За результатами перевірки 100 звітів про інспекцію, що формують за результатами агрегації результатів розпізнавання вздовж труби, отримано оцінку повної ймовірності помилкових рішень під час формування звітів, яка не перевищує 3%, що свідчить про прийнятний для практичного використання рівень точності згенерованих системою звітів.

У результаті виконаного дослідження було розроблено узагальнену та деталізовану функціональну модель процесу оптимізації параметрів системи класифікаційного аналізу даних в рамках запропонованої інформаційної технології. Розроблено узагальнену та деталізовану функціональну модель процесу аналізу даних відеоінспекції для формування звітів в рамках запропонованої інформаційної технології. Запропоновано структуру і програмну реалізацію хмарного сервісу для автоматичного формування звіту про функціональний стан труб водовідведення за даними відеоінспекції. Перевагою даного рішення є можливість асинхронної обробки множини відео за рахунок черги повідомлень та пакетного оброблення задач.

Одержані наукові результати досліджень у вигляді інформаційного та програмного забезпечення впроваджено під час підготовки персоналу та проведення інспекції каналізаційної мережі в КП "Міськводоканал" СМР (м. Суми), під час розроблення інтелектуальної системи відеомоніторингу інфраструктурних об'єктів в Товаристві з обмеженою відповідальністю "РОЗУМНІ ТЕХНОЛОГІЇ «ТИГУЛ»" (м. Суми), під час розроблення програмного забезпечення модуля машинного зору для дистанційно-керованої мобільної платформи в Товаристві з обмеженою відповідальністю "НОРД ТРЕК" (м. Шостка, Сумська область), під час розроблення системи класифікаційного аналізу даних

відеоінспекції трубопроводу водовідведення в компанії Molfar.AI sp. z o.o., (місто Гданськ, Польща), а також у навчальному процесі кафедри комп'ютерних наук Сумського державного університету під час викладання дисципліни "Introduction to Data Science".

Ключові слова: труби водовідведення, відеоінспекція, екстрактор ознак, нейронні моделі, вирішувальні правила, оптимізація, самонавчання, розпізнавання зображень, інформаційний критерій, модель відображення множин.

ABSTRACT

Zaretsky M.O. Models and methods of intelligent information technology for assessing the functional state of sewage pipes. - A qualification scientific work in the form of a manuscript. Thesis for obtaining a doctor of philosophy degree in the field of information technology, specializing in 122 - Computer Science. - Sumy State University, Sumy, 2023. The dissertation is devoted to solving the actual scientific and applied problem of automating the process of assessing the functional state of sewage pipelines based on video inspection data through the development of information technology for intelligent data analysis. The application of the developed information technology for assessing the functional state of sewage pipes allows increasing the accuracy of automatic formation of inspection reports.

An analysis of scientific sources revealed that visual observations in sewer pipe inspections differ significantly from images in open datasets, reducing the efficiency of transfer learning techniques for initial data analysis model initialization. Meanwhile, the ambiguity of encoding complex configuration defects, high variability of visual observations, the routine and costly nature of data labeling, result in limited volume and high imbalance of labeled training data, which restricts the effectiveness of traditional supervised learning methods. The encoding of the functional state of the pipe is dependent on changes in the water level, camera orientation, and defects during video inspection. The presence of a significant number of artifacts and noise necessitates consideration of spatio-temporal contextual information. Therefore, one of the complex tasks that the

dissertation aims to research is the creation of intelligent information technology for machine learning for automatic data analysis of sewer pipe video inspection under conditions of limited training data volume and resources available for system training and maintenance. At the same time, solving this task involves the necessity of analyzing spatio-temporal contextual information.

The object of the study is the process of assessing the functional state of sewer pipes. The subject of the study is models and methods of information technology for assessing the functional state of sewer pipes.

The dissertation defines the relevance of applying information technology in the field of sewer pipe inspection. For sewer pipe inspections, video cameras are used, including thermal imaging cameras, sonars, ground-penetrating radars, laser scanners, electro scanners, and their combinations with the aim of detecting defects and assessing the integrity of the pipe walls and the state of the ground around the pipes. Based on the conducted analysis of methods and approaches to sewer pipe inspection, it has been established that video cameras are the primary source of information, as video inspection requires the least overhead costs, and the informativeness of the obtained data ensures the detection of most important types of defects. The main trend in the development of modern data analysis technologies for video inspection is associated with the implementation of machine vision and artificial intelligence technologies, although these technologies do not yet provide a sufficient level of automation due to the inadequate reliability of the formed reports. Therefore, their improvement is an important direction of research. Based on the analysis of the sewer pipe video inspection process, the need for creating a new intelligent information technology for assessing the functional state of sewer pipes by using the ideas and methods of model composition, contrastive learning, noise-resistant coding, and information-extreme intelligent technology (IEI-technology) has been substantiated. Thus, the main direction for enhancing the effectiveness of models for evaluating the functional state of sewer pipes lies in adapting the input mathematical description and constructing classification decision rules during the learning process. These rules are intended for context recognition, defect recognition in individual contexts, and water level change assessment, all achieved through maximizing efficiency criteria.

For the first time, a method has been developed for training a hierarchical feature extractor, which, unlike known ones, ensures an increase in the functional efficiency of decision rules under conditions of noisy and imbalanced training data by combining the principles of knowledge transfer, siamese networks, contrastive self-learning, and self-correcting binary codes. The proposed machine learning method provides an accuracy of 98% for classification analysis of defects in the context of observations with a forward-facing camera orientation along the pipe, which is 11% higher than the result of the traditional end-to-end learning approach with a fully connected classification layer. Given the significant data imbalance present in the context of observations corresponding to the camera orientation towards the side wall, the effectiveness of defect classification is assessed using the F1-metric and amounts to $F1=0.977$, which is 5% higher than within the traditional approach. The obtained accuracy of defect classification within the proposed approach exceeds by 6% the results obtained by other authors, and allows for greater accuracy of reports on the functional state of sewer pipes.

Feature extraction models have been refined by combining local and spatio-temporal visual contextual information, which allows to increase the informativeness of the feature description of video inspections of the sewer pipeline under conditions of incomplete definiteness. The advantage of using a 1D convolutional model over LSTM and GRU networks for analyzing inter-frame dependencies has been established. Additionally, it has been empirically proven that the regular convolutional architecture of the temporal network is the most effective during the analysis of inter-frame dependencies in pipe video inspection. The micro-averaged F1-score of the feature extractor trained within the proposed approach for recognizing the context of observations from an alphabet of 11 classes is 0.92, which is 5% higher than the result for a model without considering inter-frame dependencies. For an alphabet of classes describing water levels from 0% to 55% in 5% increments, the micro-averaged F1-score within the proposed approach is 0.879 and exceeds by 9% the value for a model without considering inter-frame dependencies.

Models and algorithms for the synthesis of decision rules for the classification analysis of observations, which are based on model composition and optimization of

parameters of classification decision rules and hyperparameters of the feature extractor according to the information criterion, have been further developed. That is, models with information-extreme decision rules for recognizing context, water level, and defects are composed for assessing the functional state of the pipe. Based on the verification of 100 inspection reports formed as a result of aggregating recognition results along the pipe, an assessment of the total probability of erroneous decisions during report formation was obtained, which does not exceed 3%. This indicates an acceptable level of accuracy for practical use of the reports generated by the system.

As a result of the conducted research, a generalized and detailed functional model of the process of optimizing the parameters of the data classification analysis system was developed within the proposed information technology. A generalized and detailed functional model of the video inspection data analysis process was also developed for report generation within the proposed information technology. The structure and software implementation of a cloud service for automatic report generation about the functional state of sewage pipes based on video inspection data have been proposed. The advantage of this solution is the ability to process multiple videos asynchronously due to the message queue and batch processing of tasks.

The obtained scientific results in the form of informational and software provision have been implemented during the training of personnel and conducting of sewer network inspections at KP "Miskvodokanal" SMR (Sumy city), during the development of an intelligent video monitoring system for infrastructure objects at the Limited Liability Company "SMART TECHNOLOGIES "TITUL"" (Sumy city), during the development of software for a machine vision module for a remotely controlled mobile platform at the Limited Liability Company "NORD TREK" (Shostka city, Sumy region), during the development of a classification analysis system for sewer pipeline video inspection data at the company Molfar.AI sp. z o.o., (Gdansk, Poland), and also in the educational process of the Department of Computer Science of Sumy State University while teaching the discipline "Introduction to Data Science".

Key words: sewer pipes, video inspection, feature extractor, neural models, decision rules, optimization, self-learning, image recognition, information criterion, set mapping model.

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА

Наукові праці, в яких опубліковані основні наукові результати дисертації

Статті у фахових наукових виданнях із переліку МОН України

1. В. В. Москаленко, А. С. Москаленко, М. О. Зарецький, та А. Г. Коробов, “Модель та алгоритм навчання системи автономної навігації для малогабаритного безпілотного апарату ,” *Біоніка інтелекту*, № 91, С. 16–24, 2018.

(Особистий внесок: здобувач проаналізував існуючі алгоритми візуальної навігації).

2. В. В. Москаленко, А. С. Москаленко та М. О. Зарецький, “Модель і алгоритм навчання детектора шкідливого трафіку на основі модифікації зростаючого нейронного газу,” *Радіоелектронні І Комп’ютерні Системи*, № 3, С. 11–19, 2018, doi: 10.32620/reks.2018.3.02.

(Особистий внесок: здобувач проаналізував існуючі алгоритми синтезу інформаційно-екстремальних вирішувальних правил).

3. В.В. Москаленко, А. А. Москаленко, А.Г. Коробов, М. О. Зарецький, та В.А. Семашко, “Модель та алгоритм навчання системи детектування малорозмірних об’єктів для малогабаритних безпілотних літальних апаратів,” *Радіоелектронні І Комп’ютерні Системи*, № 4, С. 41–52, 2018, doi: 10.32620/reks.2018.4.04.

(Особистий внесок: здобувач проаналізував існуючі алгоритми екстракції ознакового опису спостережень для детектування малорозмірних об’єктів інтересу на зображенні).

4. В. В. Москаленко, М. О. Зарецький, та А. А. Москаленко, “Модель та метод навчання системи класифікації рухомих об’єктів для малогабаритного безпілотного апарату,” *Радіоелектронні І Комп’ютерні Системи*, №. 2, С. 108–117, 2019, doi: 10.32620/reks.2019.2.10.

(Особистий внесок: здобувач проаналізував існуючі алгоритми машинного навчання для розпізнавання об’єктів інтересу в процесі руху).

Статті у наукових періодичних виданнях інших держав та виданнях, які індексуються наукометричними базами даних Scopus та Web of Science

5. V. V. Moskalenko, A. S. Moskalenko, A. G. Korobov, and M. O. Zaretsky, “A Model and Training Algorithm of Small-Sized Object Detection System for a Compact Aerial Drone,” *Radio Electronics, Computer Science, Control*, vol. 0, no. 1, Apr. 2019, doi: <https://doi.org/10.15588/1607-3274-2019-1-11>. (Web of Science)

(Особистий внесок: здобувач проаналізував існуючі алгоритми детектування об’єктів інтересу на зображенні за умов ресурсних та інформаційних обмежень).

6. В. В. Москаленко, М. О. Зарецький, А. А. Москаленко, А. А. Кудрявцев, and В. Ю. Семашко, “Багатошарова модель та метод навчання для детектування шкідливого трафіку на основі ансамблю дерев рішень,” *Радіоелектронні І Комп’ютерні Системи*, № 2, С. 92–101, Apr. 2020, doi: 10.32620/reks.2020.2.08. (Scopus)

(Особистий внесок: здобувач розробив алгоритм оцінювання ефективності інформаційно-екстремальних вирішувальних правил).

7. V. V. Moskalenko, M. O. Zaretskyi, A. S. Moskalenko, and V. Lysyuk, “Sewer Pipe Defects Classification Based on Deep Convolutional Network with Information-Extreme Error-Correction Decision Rules,” in *Springer eBooks*, 2020, pp. 253–263. doi: 10.1007/978-3-030-61656-4_16. (Scopus)

(Особистий внесок: здобувач розробив модель та алгоритм навчання для розпізнавання дефектів труб водовідведення, коли камера орієнтована вперед вздовж труби, а також виконав експериментальні дослідження).

8. В. В. Москаленко, М. О. Зарецький, Я. Ю. Ковальський, та С. С. Мартиненко, “Модель і метод навчання класифікатора контекстів спостереження на зображеннях відеоінспекції стічних труб,” *Радіоелектронні І Комп’ютерні Системи*, № 3, С. 59–66, 2020, doi: 10.32620/reks.2020.3.06. (Scopus)

(Особистий внесок: здобувач сформував вхідний математичний опис та розробив глибоку модель з інформаційно-екстремальними вирішувальними

правилами для розпізнавання контексту спостереження на кадри відеоінспекції, а також виконав експериментальні дослідження).

9. В. В. Москаленко, М. О. Зарецький, А. С. Москаленко, А. Г. Коробов, та Я. Ю. Ковальський, “Багатоетапний метод глибинного навчання з попереднім самонавчанням для класифікаційного аналізу дефектів стічних труб,” *Радіоелектронні І Комп’ютерні Системи*, № 4, С. 71–81, 2021, doi: 10.32620/reks.2021.4.06. (Scopus)

(Особистий внесок: здобувач розробив багатоетапний метод навчання глибокої моделі з ініціалізацією вагових коефіцієнтів в процесі контрастного самонавчання для розпізнавання дефектів труб водовідведення, коли камера орієнтована вбік до стінки труби).

10. В. В. Москаленко, М. О. Зарецький, А. Г. Коробов, Я. Ю. Ковальський, А. Ф. Шаєхов, В. А. Семашко, та ін., “Модель та метод навчання для класифікаційного аналізу рівня води в стічних трубах за даними відео інспекції,” *Радіоелектронні І Комп’ютерні Системи*, № 2, С. 4–15, 2021, doi: 10.32620/reks.2021.2.01. (Scopus)

(Особистий внесок: здобувач сформував вхідний математичний опис, розробив глибоку модель для розпізнавання рівня води в трубі на зображеннях відео інспекції з урахуванням міжкадрових залежностей, а також виконав експериментальні дослідження).

11. V. V. Moskalenko, M. O. Zaretsky, A. S. Moskalenko, A. O. Panych, and V. V. Lysyuk, “A Model and Training Method for Context Classification in Cctv Sewer Inspection Video Frames,” *Radio Electronics, Computer Science, Control*, no. 3, pp. 97–108, Oct. 2021, doi: <https://doi.org/10.15588/1607-3274-2021-3-9>. (Web of Science)

(Особистий внесок: здобувач сформував вхідний математичний опис, розробив глибоку модель для розпізнавання контексту спостереження на зображеннях відео інспекції з урахуванням міжкадрових залежностей, а також виконав експериментальні дослідження).

12. V. V. Moskalenko, A. S. Moskalenko, A. G. Korobov, and M. O. Zaretsky, “Image Classifier Resilient to Adversarial Attacks, Fault Injections and Concept Drift –

Model Architecture and Training Algorithm,” *Radio Electronics, Computer Science, Control*, no. 3, p. 86, Oct. 2022, doi: 10.15588/1607-3274-2022-3-9. (Web of Science)

(Особистий внесок: програмно реалізував згорткову нейронну мережу з інформаційним пляшковим горлом та гіперсферичними контейнерами класів розпізнавання).

Наукові праці, які засвідчують апробацію матеріалів дисертації

13. В. В. Москаленко, М. О. Зарецький, Д. О. Лєвокутник, “Модель і метод навчання детектора малорозмірних об’єктів за умов ресурсних та інформаційних обмежень,” *XIV International scientific conference «Intellectual systems of decision-making and problems of computational intelligence»*. с. Залізний Порт, 21-25 трав. 2019 С. 116-118

(Особистий внесок: здобувач проаналізував існуючі методи машинного навчання глибоких моделей детектування об’єктів).

14. V. V. Moskalenko, A. S. Moskalenko, and M. O. Zaretskyi, “Model and training method of moving object classification system for a compact drone,” *14th International Conference on Computer Sciences and Information Technologies (CSIT)*, 2019, pp. 56–61. doi: 10.1109/stc-csit.2019.8929795. (Scopus & Web of Science)

(Особистий внесок: здобувач проаналізував існуючі алгоритми детектування об’єктів інтересу на місцевості в процесі руху).

15. M. Zaretskyi V. Moskalenko, A. Moskalenko, and A. Shaiekhov, “Multi-Layer Model and Training Method for Information-Extreme Malware Traffic Detector,” *The Third International Workshop on Computer Modeling and Intelligent Systems*, April 27-May 1, 2020, pp. 288-299 Available at: <https://ceur-ws.org/Vol-2608/paper22.pdf>. (accessed on May 1 2021). (Scopus)

(Особистий внесок: здобувач запропонував метод обчислення еталонного вектора кожного класу розпізнавання в двійковому просторі для синтезу інформаційно-екстремальних вирішувальних правил).

16. V. V. Moskalenko, A.S. Moskalenko, M.O. Zaretskyi, and V. Lysyuk, “Deep Feature Extractor with Information-Extreme Decision Rules for Visual Classification of Sewer Pipe Defects and its Training Method,” *Data Stream Mining & Processing*, pp. 191–195, 2022 doi: 10.1109/dsmp47368.2020.9204083. (Scopus & Web of Science)

(Особистий внесок: здобувач сформував вхідний математичний опис та розробив глибоку модель з інформаційно-екстремальними вирішувальними правилами для розпізнавання дефектів у випадку, коли камера орієнтована вперед вздовж труби).

17. М. О. Зарецький, “Нейромережеве вимірювання рівня води в трубопроводі водовідведення за даними відеоінспекції,” *Міжнародна наукова конференція молодих учених Інформатика, Математика, Автоматика, ІМА-2023*. Суми, Україна - Астана, Казахстан, 24–28 квітня 2023, С. 45-46. Available: https://essuir.sumdu.edu.ua/bitstream-download/123456789/91550/1/Conf_IMA_2023.pdf . (accessed on: May 22, 2023).

(Особистий внесок: здобувач проаналізував існуючі методи машинного навчання глибоких моделей детектування об’єктів).

Наукові праці, які додатково відображають наукові результати дисертації

18. В.В. Москаленко, А.С. Москаленко, М.О. Зарецький, В.В. Нагорний, Свідоцтво про реєстрацію авторського права на твір № 89337, комп’ютерна програма «Інформаційно-екстремальний класифікатор багатовимірних спостережень з метаевристичною оптимізацією параметрів функціонування» заяв. 2019-04-05; опубл. 2019-06-05.

(Особистий внесок: програмна реалізація фітнес-функції мета-евристичного алгоритму, в якій обчислюється інформаційний критерій ефективності моделі аналізу даних на тестовій вибірці)

19. В.В. Москаленко, А.С. Москаленко, М.О. Зарецький, А.О. Панич, Свідоцтво про реєстрацію авторського права на твір № 106341, комп’ютерна

програма «Програма для розпізнавання та видалення текстових анотацій з відеозображень» заяв. 2021-07-05; опубл. 2021-07-16.

(Особистий внесок: програмна реалізація функції заповнення виділеної ділянки зображення на основі інформації про оточуючі ділянки зображення)

20. В.В. Москаленко, А.С. Москаленко, М.О. Зарецький, А.О. Панич, А.Г. Коробов, Свідоцтво про реєстрацію авторського права на твір № 107201, комп'ютерна програма «Реалізація алгоритму інформаційно-екстремального навчання глибокої моделі аналізу даних з використанням принципів сіамських мереж та завадозахищеного кодування» дата реєстр. 11.08.2021.

(Особистий внесок: здобувач виконав програмну реалізацію алгоритму навчання глибокої згорткової моделі для побудови інформаційно-екстремального класифікатора зображень відеоінспекції стічних труб).

21. В.В. Москаленко, А.С. Москаленко, М.О. Зарецький, А.Г. Коробов, Свідоцтво про реєстрацію авторського права на твір № 113206 «Програма машинного навчання класифікатора зображень з підвищеною робастністю до шуму і змагальних атак» заяв. 2022-06-06; опубл. 2022-07-29.

(Особистий внесок: програмна реалізація алгоритму аугментації даних з використанням шумових шаблонів)

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, СКОРОЧЕНЬ І ТЕРМІНІВ ..	19
ВСТУП.....	20
РОЗДІЛ 1 АНАЛІЗ СУЧАСНИХ НАПРЯМКІВ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ ПРОЦЕСУ ІНСПЕКЦІЇ ТРУБ ВОДОВІДВЕДЕННЯ.....	26
1.1 Сучасний стан та тенденція розвитку технологій інспекції труб водовідведення.....	26
1.2 Методи, моделі та алгоритми екстракції ознакового опису відеозображень	38
1.3 Методи, моделі та алгоритми ідентифікації об’єктів інтересу на зображеннях відеоінспекції	51
Висновки до першого розділу	62
РОЗДІЛ 2 РОЗРОБЛЕННЯ МОДЕЛІ ТА МЕТОДУ НАВЧАННЯ ДЛЯ ОЦІНЮВАННЯ ФУНКЦІОНАЛЬНОГО СТАНУ ТРУБ ВОДОВІДВЕДЕННЯ....	64
2.1 Аналіз слабо формалізованого процесу оцінювання функціонального стану труб водовідведення та обґрунтування вибору методу досліджень	64
2.2 Формалізована постановка задачі	72
2.3 Моделі відображення множин для опису процесу навчання.....	75
2.4 Критерії оцінювання ефективності моделі класифікаційного аналізу даних	80
2.5 Моделі екстракції ознакового опису спостережень.....	87
Висновки до другого розділу.....	92
РОЗДІЛ 3 МЕТОДИ ОПТИМІЗАЦІЇ ПАРАМЕТРІВ МОДЕЛІ АНАЛІЗУ ДАНИХ ВІДЕОІНСПЕКЦІЇ ТРУБ ВОДОВІДВЕДЕННЯ.....	96
3.1 Метод оптимізації параметрів моделі класифікаційного аналізу зображень з використання нерозмічених та розмічених даних.....	96
3.2 Оптимізація параметрів моделі розпізнавання контекстів спостереження .	102
3.3 Оптимізація параметрів моделі оцінки рівня води в трубопроводі водовідведення	111
3.4 Оптимізація параметрів моделей розпізнавання дефектів труб водовідведення	119

Висновки до третього розділу	130
РОЗДІЛ 4 ІНФОРМАЦІЙНЕ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ ОЦІНЮВАННЯ ФУНКЦІОНАЛЬНОГО СТАНУ ТРУБОПРОВОДУ ВОДОВІДВЕДЕННЯ	134
4.1 Структура та функції інтелектуальної системи оцінювання функціонального стану труб водовідведення за даними відеоінспекції	134
4.2 Формування звітів про функціональний стан трубопроводу водовідведення	139
4.3 Програмний інструментарій та реалізація інтелектуальної системи оцінювання функціонального стану труб водовідведення.....	145
Висновки до четвертого розділу	155
ВИСНОВКИ.....	157
ДОДАТКИ.....	174
ДОДАТОК А Акти впровадження результатів дисертаційної роботи.....	174
ДОДАТОК Б Список публікацій за темою дисертації.....	179

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, СКОРОЧЕНЬ І ТЕРМІНІВ

ІЕІ-технологія – інформаційно-екстремальна інтелектуальна технологія.

CNN – Convolutional Neural Network.

LSTM – Long Short-Term Memory.

GRU – Gated Recurrent Units.

PACP – Pipeline Assessment Certification Program.

YOLO – You only look once.

SSD – Single Shot Detector.

HTTP – Hypertext Transfer Protocol

REST – Representational State Transfer.

Amazon S3 – Amazon Simple Storage Service

ВСТУП

Актуальність теми. Застосування інформаційних технологій відеоінспекції забезпечує найкраще співвідношення ціни та інформативності щодо оцінювання функціонального стану інфраструктури водопостачання і водовідведення. Процес аналізу відеоданих інспекції є трудомістким і рутинним, вимагає уважності і точного дотримання стандартів кодування, однак результат не захищений від суб'єктивних факторів оператора. Тому в задачі аналізу даних відеоінспекції трубопроводу водовідведення для виявлення дефектів і їх кодування згідно існуючих стандартів є потреба в підвищенні рівня автоматизації і зниження когнітивного навантаження на оператора. Традиційним підходом до автоматизації аналізу даних є використання інтелектуальних інформаційних технологій, які моделюють когнітивні процеси, що притаманні людині під час прийняття рішень. Дослідження, пов'язані з розробленням інтелектуальних технологій аналізу відео зображень різного призначення, в науково-технічній літературі досить широко висвітлені завдяки ідеям та науковим здобуткам передусім, А. С. Довбиша, С. О. Суботіна, О. А. Бодянського, М. І. Шлезінгера, Я. Лекуна, Й. Бенджіо, Д. Хінтона, Р. Сігварта, Д. Скарамuzzi, Х. Ньето, А. Котуса та інших вчених і їх науковими школами. При цьому питання підвищення функціональної ефективності за умов ресурсних та інформаційних обмежень усе ще залишається недостатньо дослідженим через науково-методологічні ускладнення, пов'язані з неповною визначеністю даних.

Одним з перспективних шляхів підвищення ефективності інформаційних технологій інтелектуального аналізу даних відеоінспекції трубопроводу водовідведення за умов ресурсних та інформаційних обмежень є використання принципів переносу знань (transfer learning) та алгоритмів настроювання гіперпараметрів. Однак візуальні спостереження в інспекції трубопроводу водовідведення досить сильно відрізняються від зображень відкритих наборів даних, на яких навчаються state-of-the-art моделі, що досить сильно знижує

ефективність техніки переносу знань. Крім того дані відео інспекції труб містять значну кількість артефактів та шумів, наприклад, засвічування поверхні, бризки води, тимчасові неполадки відеопристрою, дефекти зображень, схожість бруду та піни на реальні дефекти. При цьому неоднозначність кодування дефектів складної конфігурації, висока варіативність візуальних спостережень, рутинність і дороговартісність розмітки даних, обумовлюють обмеженість обсягу і високу незбалансованість навчальних даних. Це призводить до збільшення впливу вибору архітектури, гіперпараметрів та початкової ініціалізації вагових коефіцієнтів на ефективність навчання. Тому одним із складних завдань, на дослідження якого спрямована дисертаційна робота, є створення інтелектуальної інформаційної технології для автоматичного оцінювання функціонального стану труб водовідведення за даними відеоінспекції за умов обмеженого обсягу навчальних даних та ресурсів, доступних для навчання та обслуговування системи. Водночас вирішення цього завдання полягає в необхідності аналізу просторово-часової інформації з урахуванням незбалансованості вхідних даних та впливу довільних початкових умов процесу аналізу даних, обумовлених змінами орієнтації камери та інших контекстних умов спостереження.

Мета і завдання дослідження відповідно до предмета та об'єкта дослідження.

Об'єкт дослідження – процес оцінювання функціонального стану труб водовідведення.

Предмет дослідження – моделі й методи інформаційної технології оцінювання функціонального стану труб водовідведення.

Метою дослідження є підвищення ефективності оцінювання функціонального стану труб водовідведення за умов неповної визначеності шляхом створення інформаційної технології машинного навчання.

Для досягнення поставленої мети необхідно вирішити такі задачі:

– проаналізувати сучасний стан і тенденцію розвитку інформаційних технологій оцінювання функціонального стану трубопроводу водовідведення;

- розробити модель екстрактора ознакового опису для врахування просторово-часового контексту відеоспостережень, що функціонує за умов неповної визначеності;
- розробити моделі розпізнавання дефектів труб водовідведення;
- розробити модель розпізнавання рівня води в трубопроводі водовідведення;
- розробити метод машинного навчання для формування інформативного ознакового опису спостережень за умов обмеженого і незбалансованого обсягу розмічених зразків;
- розробити метод побудови вирішувальних правил системи відеоінспекції трубопроводу водовідведення;
- розробити інформаційну технологію синтезу інтелектуальної системи відеоінспекції трубопроводу водовідведення;
- впровадити результати дисертаційного дослідження в практику створення інформаційних систем відеоінспекції інфраструктурних об'єктів водовідведення.

Гіпотеза дослідження. Побудова композиції моделей аналізу даних з урахуванням просторово-часового контексту та використання багатоетапного методу навчання на нерозмічених та розмічених даних можуть забезпечити підвищення ефективності оцінювання функціонального стану трубопроводу за даними відеоінспекції.

Методи дослідження базуються на принципах і методах технології штучних нейронних мереж, методах декомпозиції, теорії інформації і кодування, методах теорії оптимізації, регуляризації та контрастного навчання, а також методах валідації моделей аналізу даних.

Наукова новизна отриманих результатів полягає в розробленні інтелектуальної інформаційної технології оцінювання функціонального стану труб водовідведення, яка забезпечує підвищення точнісних характеристик оцінювання функціонального стану труб водовідведення за даними відеоінспекції.

У результаті проведеного дослідження здобувачем одержано такі наукові результати:

вперше розроблено:

1) метод навчання ієрархічного екстрактора ознак, що на відміну від відомих, забезпечує підвищення функціональної ефективності вирішувальних правил за умов зашумленості та незбалансованості навчальних даних шляхом поєднання принципів переносу знань, сіамських мереж, контрастного самонавчання та самокоректуючих двійкових кодів;

удосконалено:

2) моделі екстракції ознакового опису спостережень шляхом поєднання локальної і просторово-часової візуальної контекстної інформації, що дозволяє підвищити інформативність ознакового опису відеоспостережень інспекції трубопроводу водовідведення за умов неповної визначеності;

одержала подальшого розвитку:

3) моделі і алгоритми синтезу вирішувальних правил для класифікаційного аналізу спостережень, які ґрунтуються на композиції моделей та оптимізації параметрів класифікаційних вирішувальних правил та гіперпараметрів екстрактора ознак за інформаційним критерієм.

Особистий внесок здобувача. Положення і результати, винесені на захист дисертаційної роботи, отримані здобувачем особисто. Серед них: самостійно виконаний аналіз існуючих алгоритмів екстракції ознак, навчання та екзамену для детектування об'єктів інтересу на зображенні, в тому числі в умовах руху; аналіз існуючих алгоритмів візуальної навігації, аналіз існуючих алгоритмів синтезу та оцінки ефективності інформаційно-екстремальних вирішувальних правил; розроблено багатоетапний метод навчання глибокої згорткової моделі для побудови інформаційно-екстремального класифікатора зображень відеоінспекції труб водовідведення; розроблено алгоритм синтезу інформаційно-екстремальних вирішувальних правил для згорткового екстрактора ознак; сформовано вхідний математичний опис для задачі розпізнавання контексту спостережень, рівня води та детектування дефектів труб водовідведення; розроблено моделі для детектора дефектів та контексту спостереження на кадрі відео інспекції труб водовідведення; розроблено моделі розпізнавання контексту та рівня води в трубах водовідведення

з урахуванням міжкадрових залежностей; розроблено програмне та інформаційне забезпечення системи аналізу даних інспекції труб водовідведення.

Апробація результатів дисертації. Наукові результати роботи доповідалися також на: 15-й Міжнародній науковій конференції «Intellectual systems of decision-making and problems of computational intelligence, ISDMCI-2019» (м. Херсон, Україна, 2019 р.), 14-й Міжнародній науковій конференції «Computer Sciences and Information Technologies, CSIT-2019» (м. Львів, Україна, 2019 р.), 3-й Міжнародній науковій конференції «Data stream mining & processing, DSMP-2020» (м. Львів, Україна, 2020 р.), 3-й Міжнародний воркшоп «Computer Modeling and Intelligent Systems, CMIS-2020» (м. Запоріжжя, Україна, 2020), Міжнародній науково-практичній конференції молодих вчених «Інформатика, Математика, Автоматика, ІМА-2023» (м. Суми, Україна-м. Астана, Казахстан, 2023 р.).

Структура та обсяг дисертації. Дисертаційна робота складається з анотації, змісту, переліку умовних скорочень, вступу, чотирьох розділів, висновку, списку використаних джерел та додатків. Повний обсяг роботи становить 184 сторінки друкованого тексту, з них зміст – на 2 стор., перелік умовних скорочень – на 1 стор., основний текст – на 140 стор., список зі 117 використаних джерел – на 14 стор., додатки – на 11 стор. Дисертація містить 68 рисунків та 8 таблиць.

Зв'язок роботи з науковими програмами, планами, темами.

Дисертаційне дослідження відповідає пріоритетним напрямкам наукових досліджень Сумського державного університету, зокрема держбюджетним науково-дослідним роботам «Інтелектуальна автономна бортова система безпілотного літального апарату для ідентифікації об'єктів на місцевості» (ДР № 0117U003934), «Методи та математичні моделі сучасних інформаційно-комунікаційних технологій» (ДР № 0112U006083), «Інформаційна технологія забезпечення резильєнтності систем штучного інтелекту для захисту кіберфізичних систем» (ДР № 0122U000782).

Практичне значення отриманих результатів. Одержані наукові результати досліджень у вигляді інформаційного та програмного забезпечення впроваджено під час підготовки персоналу та проведення інспекції каналізаційної мережі в КП

"Міськводоканал" СМР (м. Суми), під час розроблення інтелектуальної системи відеомоніторингу інфраструктурних об'єктів в Товаристві з обмеженою відповідальністю "РОЗУМНІ ТЕХНОЛОГІЇ «ТИТУЛ»" (м. Суми), під час розроблення програмного забезпечення модуля машинного зору для дистанційно-керованої мобільної платформи в Товаристві з обмеженою відповідальністю "НОРД ТРЕК" (м. Шостка, Сумська область), під час розроблення системи класифікаційного аналізу даних відеоінспекції трубопроводу водовідведення в компанії Molfar.AI sp. z o.o., (місто Гданськ, Польща), а також у навчальному процесі кафедри комп'ютерних наук Сумського державного університету під час викладання дисципліни "Introduction to Data Science".

Подяка. Автор роботи щиро дякує науковому керівнику, к.т.н., доценту, доценту кафедри комп'ютерних наук Сумського державного університету Москаленку В'ячеславу Васильовичу за корисні поради та зауваження, підтримку і віру під час підготовки даної дисертаційної роботи.

РОЗДІЛ 1

АНАЛІЗ СУЧАСНИХ НАПРЯМКІВ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ ПРОЦЕСУ ІНСПЕКЦІЇ ТРУБ ВОДОВІДВЕДЕННЯ

1.1 Сучасний стан та тенденція розвитку технологій інспекції труб водовідведення

Мережі каналізаційних трубопроводів стали основною проблемою сучасних муніципалітетів у всьому світі, оскільки ці мережі занадто старі і досягають свого проектного терміну експлуатації; тим часом зростаючі вимоги та обмежений бюджет ускладнюють вирішення проблеми. Через погіршення стану каналізаційних мереж протягом терміну їх служби, навіть якщо вони сьогодні перебувають у дуже хорошому стані, вони все одно потребуватимуть все більших інвестицій із часом внаслідок старіння. Крім того, трубопроводи водовідведення схильні до засмічення аж до повної втрати здатності виконувати свої функції [1]. Для вчасного усунення засмічення та запобігання подальшим серйозним і дорогим пошкодженням, стан каналізаційних систем потрібно контролювати за допомогою своєчасної та всебічної періодичної оцінки.

Підземну інфраструктуру складно інспектувати, оскільки вона прихована під землею. Ця ж обставина робить труби водовідведення дорогими для ремонту та заміни. Безпосередній огляд людиною, на відміну від інших інфраструктур, таких як дорожні покриття та тунелі, просто недоцільний для підземних інфраструктур через велику множину заглиблених трубопроводів і мереж, небезпечне та шкідливе для здоров'я середовище каналізації, низьку видимість та малий розмір труб. У відповідності до програми сертифікації щодо оцінки трубопроводів (PACP - Pipeline Assessment Certification Program) у каналізаційній системі є широкий спектр дефектів, і виявлення всіх їх за допомогою звичайних методів може бути нездійсненним [2]. Дефекти трубопроводів каналізації можна класифікувати на дефекти структурні, конструктивні і операційні та дефекти технічного обслуговування (рис. 1.1).

За стандартом РАСР [3] класи дефектів додатково розділяються на точкові (Point Defects) та протяжні (Continuous Defects). Перший включає дефекти, що кодуються індивідуально, по мірі їх виникнення. Другий – пошкодження, які проходять уздовж труби, наприклад, коріння. Протяжні дефекти класифікуються на дійсні (Truly Continuous Defect) – ті що проходять уздовж труби безперервно, упродовж більше одного метра, наприклад, подовжня тріщина; та повторювані (Repeated Continuous Defect) – дефекти чи спостереження, які виникають принаймні в 3 з 4 з'єднань (75%) уздовж труби. Корені, що з'являються на більшості стиків труб, є прикладом протяжних дефектів, що повторюються.

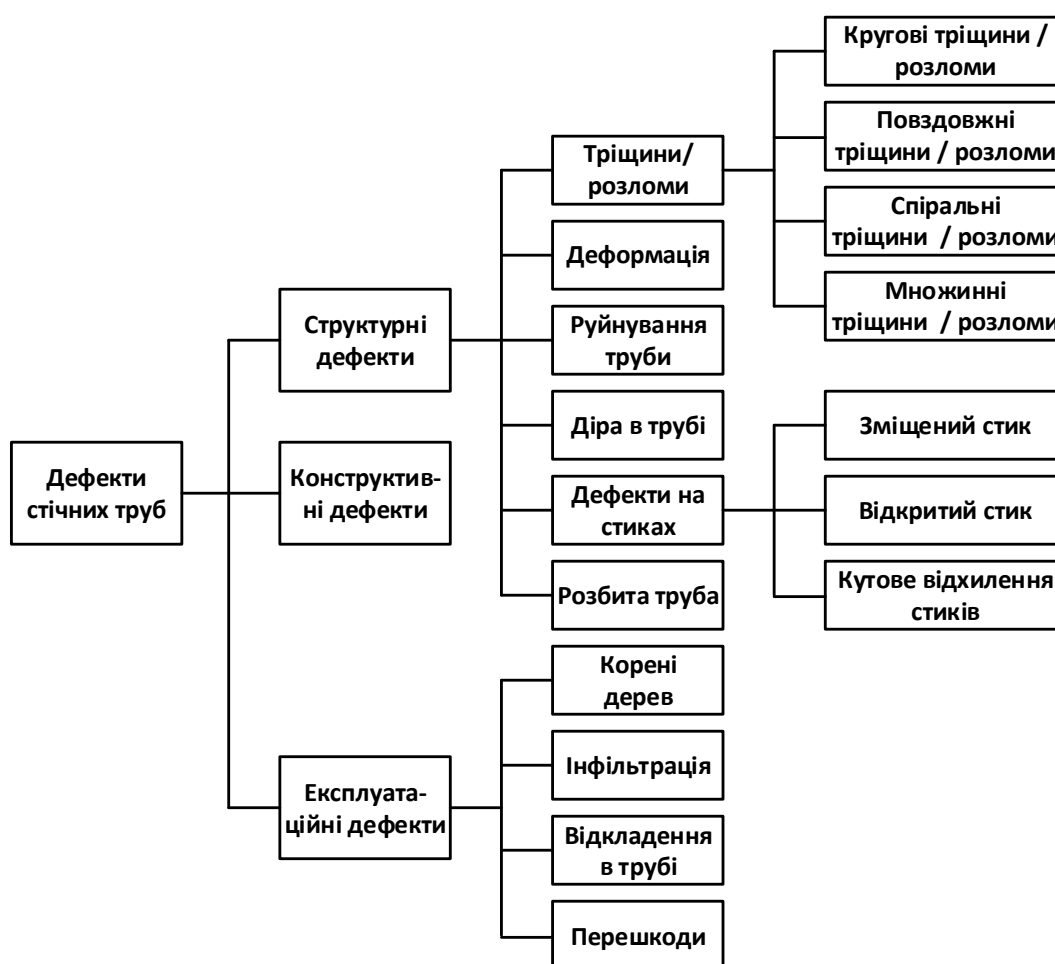


Рисунок 1.1 – Основні типи дефектів труб водовідведення згідно стандарту РАСР

Місця з'єднання труб (joint) – є одними з найбільш слабких складових частин каналізаційної системи, дефекти яких не тільки зустрічаються досить часто, а й

можуть призвести до повної втрати трубою водовідведення здатності виконувати свою цільову функцію. Крім того, будь-яке пошкодження, що виникає на відстані 8 дюймів або 200 мм від стику безперервних сегментів труб додатково помічається в звіті інспекції [3].

Втрата гідравлічної пропускної здатності зумовлена просіданням труби (рис. 1.2).

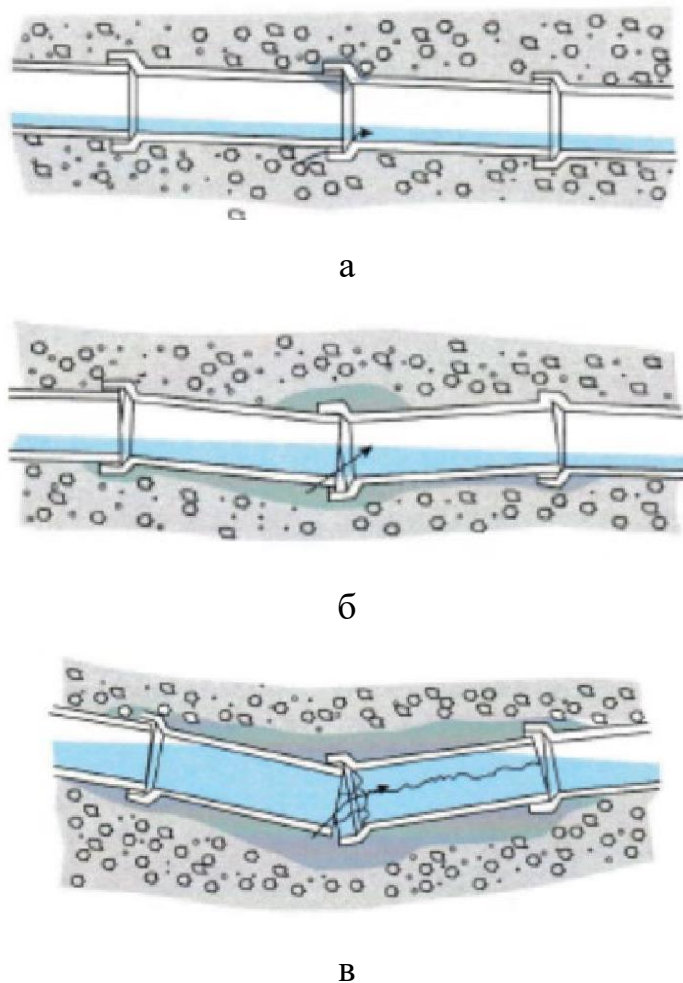


Рисунок 1.2 – Ілюстрація етапів просідання труби [3]: а – утворення розриву; б – інфільтрація; в – руйнування труби

Просідання труби відбувається в три етапи:

- утворення розриву, спричиненого можливими дефектами стику труб;

- інфільтрація ґрунтових вод з утворенням порожнин навколо стиків і, як наслідок, втрата підтримки ґрунту навколо труби дозволяє їй рухатись та просідати;

- нерівномірне навантаження на труби через зміщення стиків і відсутність підтримки з боку навколишнього ґрунту призводить до розтріскування або руйнування труб.

В міру зростання навантаження тріщини або розломи можуть перетворюватися на руйнування та деформації. Камера може бути затоплена через підвищення рівня води. Поширеною проблемою є неможливість повноцінного завершення інспекції через обмеження обладнання або перешкоди у вигляді зруйнованого сегменту труби. На рис. 1.3 показано як зміщення стику труб може стати причиною ускладнення обстеження. В трубі між двома перешкодами є тріщини, розломи і уламки, але, оскільки ця частина не могла бути обстежена, ці дефекти не будуть задокументовані належним чином [3].

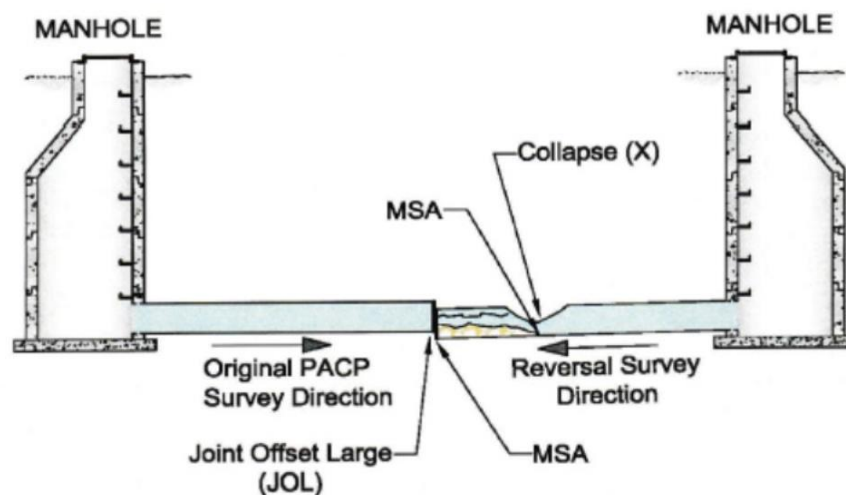


Рисунок 1.3 – Ілюстрація прикладу впливу зміщення стику труб водовідведення на ускладнення оцінки проведеної інспекції [3]

Вищезазначені проблеми завжди були ключовими атрибутами, що стимулюють розробку більш складних інструментів інспекції каналізаційних трубопроводів. Розвиток сенсорних та оптичних технологій спростив та

пришвидшив впровадження та вдосконалення нових методів детектування дефектів.

Методи інспекції каналізаційних трубопроводів можуть бути розділені на чотири групи наступним чином. Перший – це технології, що базуються на технічному зорі, де використовується відеокамера і повністю залежать від того, на що вона здатна. Другий – це структурні та застиляючі інструменти, що використовуються для перевірки цілісності стінок труб та стану ґрунту навколо труб. Третій – це специфічні технології для окремих видів дефектів, і нарешті, гібридні технології, які є комбінацією двох або більше інструментів (рис. 1.4) [2,4].



Рисунок 1.4 – Види інструментів інспекції трубопроводів

Одним з найпростіших методів інспекції труб водовідведення є використання традиційних відеокамер для запису відео в підземних трубопроводах. Вони застосовуються для огляду труб, що можуть бути занадто малими або небезпечними для людей. На ранніх стадіях камери відеоспостереження були закріплені на тросі між двома люками. З часом камери відеоспостереження почали встановлюватись на гусеничному шасі або поплавку [5]. Оператори можуть контролювати рух робота, як і камери, з далеких відстаней. Камера реєструє внутрішню поверхню, стан труб та надає інформацію вище рівня стічних вод. Пізніше експерти використовують записане відео для інтерпретації, коментування та прийняття рішень стосовно стану труб. На рис. 1.5 можна побачити кадр даного дослідження.



Рисунок 1.5 – Кадр відеоінспекції труби водовідведення

Одним з удосконалень технології інспекції каналізації є лазерний профайлер. Лазерний профайлер – це технологія, здатна виявляти і кількісно визначати зміни у вертикальній і горизонтальній формі трубопроводів. Існує два типи лазерних профайлерів : двовимірний (2-D) лазерний профайлер та тривимірний (3-D) лазерний профайлер [6]. Двовимірна технологія лазерного профайлера базується на кільці світла, генерованому лазером, навколо стінки труб. Камера

відеоспостереження, яка кріпиться на той самий сканер, виявляє кільце світла і зберігає лазерне зображення для подальшого аналізу. Використовуючи лише відеоспостереження, оператор може не спостерігати жодного прогину вздовж труб аналізуючи записане відео. Однак використання двовимірного лазерного профайлера чітко представляло б реальний стан труб [4]. Тривимірний лазерний профайлер є більш досконалою системою. Він використовує лазерну точку та пучки, які мають приймач і двосторонній передавач. Вихідні дані інспекції – це тривимірна діаграма координат X, Y та Z труб (рис. 1.6).

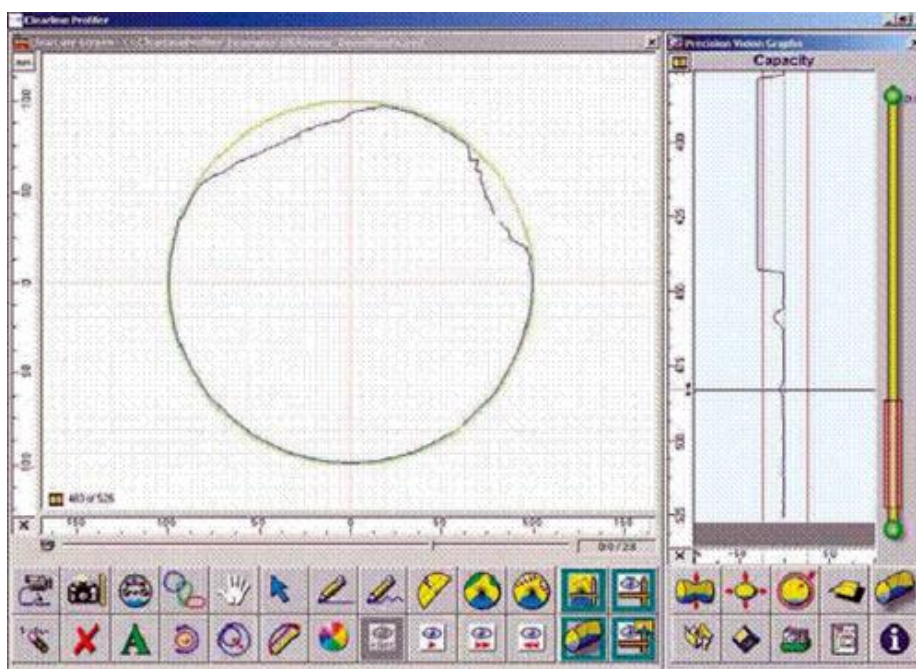


Рисунок 1.6 – Результат дослідження лазерним профайлером

Лазер здатний фіксувати будь-які незначні зміни в геометрії труб, які важко виявити за допомогою камери відеоспостереження. З іншого боку, камери відеоспостереження можуть виявляти тріщини та переломи. Використання обох технологій забезпечує економію коштів та більш досконалі плани відновлення труб водовідведення [4].

Для діагностики труб водовідведення також використовують акустичні технології [7]. Наприклад, сонар може сканувати поверхню за рахунок випромінювання звукової енергії з частотою, вищою ніж може почути людина, і

здійснювати аналіз результатів відбивання хвиль. Звукові хвилі проходять крізь внутрішню частину труби. Хвилі відбиваються щоразу, коли відбувається зміна щільності матеріалу. Частина відбитих хвиль проходить через нове середовище, тоді як інші повертаються на поверхню. Вибір частоти хвилі значно впливає на зображення, створене сонаром (рис. 1.7). При збільшенні акустичної частоти проникаюча сила зменшується. Крім того, швидкість руху також може впливати на якість зображення. Рухаючись зі швидкістю 100 мм в секунду пристрій може виявляти критичні дефекти, але може не виявляти дрібні дефекти [5].

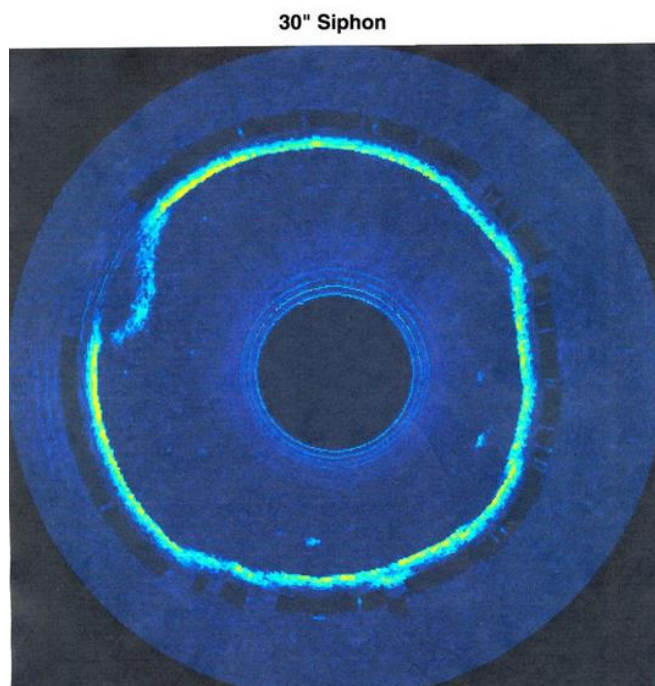


Рисунок 1.7 – Результат дослідження сонарим профайлером

Багато галузей промисловості, включаючи медичну, аерокосмічну та нафтогазову, використовують сонарну технологію [7]. Промисловість інспекції каналізації може застосовувати сонари з будь-яким типом матеріалів трубопроводів і його використовують найчастіше для кількісної оцінки осілих відкладень, таких як жир та сміття.

У складних роботах сонар використовується з лазерними профайлерами та камерами відеоспостереження. Датчик ехолота в основному використовується нижче лінії стічних вод для вимірювання обсягу осілих відкладень. Сонар може

надати експертам загальний обсяг осілих відкладень та відсоток завалу труби. При використанні датчика гідролокатора з лазерними профілями та камерами відеоспостереження генерується 3-D модель для відображення існуючого стану труб [8].

Кілька зображень сонарів дають реальний стан труб нижче лінії потоку, тоді як лазерні профайлери та камери відеоспостереження надаватимуть інформацію над лінією потоку. Інтеграція лазерного профайлер та сонара надають інформацію про геометричну форму труб, демонструючи будь-яку деформацію, руйнування стінки внаслідок корозії та осідання відкладень на дні труб [7, 9].

Для максимальної інформативності спостереження деякі виробники обладнання для інспекції труб водовідведення застосовують методи цифрового сканування на основі відеокамер з кутом огляду 360 градусів (так зване риб'яче око) [9, 10]. Камери встановлені на гусеничному ході, який рухається вздовж труб. Для цифрового сканування використовуються два типи камер високої роздільної здатності. Ці дві камери забезпечують інформацію про частини труб, а також круговий вигляд труб (рис. 1.8).

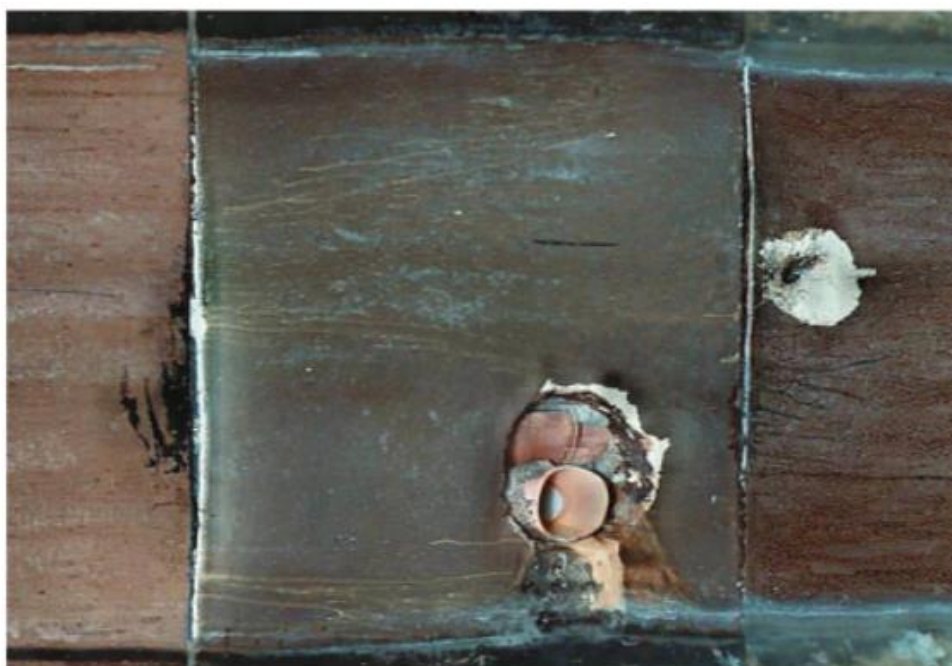


Рисунок 1.8 – Результат цифрового сканування бічних стінок труби

Швидкість інспекції і перевірки результатів для методу цифрового сканування може бути в два-три рази більшою, ніж звичайна відеоінспекція. Крім того, оператор має панорамний вид труб, оскільки цифрове сканування пропонує цю інформацію автоматично.

Ще одним джерелом інформації про функціональний стан труби водовідведення є інфрачервоне випромінювання [11]. Інфрачервоне випромінювання тече з теплих у прохолодні місця. Цей принцип застосовується до будь-якого матеріалу. Однак кожен матеріал по-різному утримує тепло, оскільки кожен має різні ізоляційні властивості. Ця концепція застосовується для перевірки каналізації. Інфрачервоне сканування використовує камеру для вимірювання інфрачервоної частоти випромінювання поверхні труб (рис. 1.9).

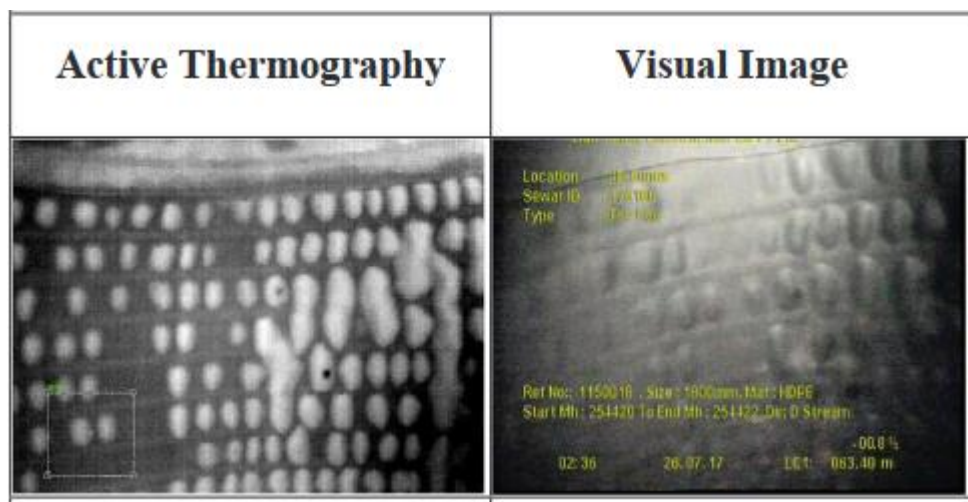


Рисунок 1.9 – Результат цифрового сканування бічних стінок труби

Обладнання для інспекції містить сенсор інфрачервоного випромінювання, оптичну головку, мікропроцесор, монітор, пристрій для збору даних, обладнання для аналізу зображень та приймачі [10, 11]. Даний пристрій є чутливий до місць витоку тепла, одна існує потреба у врахуванні багатьох зовнішніх факторів, таких як тип і вологість ґрунту, вологість та температура повітря, швидкість вітру і тому подібне.

Для отримання більш глибокої інформації деякі розробники систем відеоінспекції використовують радар підповерхневого зондування (георадар) [12].

Принцип дії такого радару ґрунтується на теорії радіолокаційних систем, де антена виробляє високочастотні радіо хвилі. Радар застосовується в трубі, тому сигнал буде проникати через стінку труби до навколишнього ґрунту. Система може працювати з використанням двох або трьох антен, які здатні виявляти різні частоти для оцінки оточення та структури самої труби. Робот SewerVUE, який застосовує концепцію георадарів, може надати інформацію про товщину стінок, товщину утеплювача, люки та стан підкладок для неметалевих труб. Робот також оснащений технологіями відеоспостереження та лазерним сканером (LIDAR) [13].

В останні роки з'явилися нові методи перевірки шляхом інтеграції двох або більше технологій для виявлення різних типів дефектів в каналізаційних системах. Ця стратегія може призвести до компенсації обмежень інших технологій, таких як технічний зір. В технології сканування і оцінювання (SSET) використовується об'єктив камери «риб'яче око» з оптичним сканером і технологією гіроскопа для забезпечення огляду всієї поверхні труб. Відскановані зображення оцифрують для генерації комп'ютерних зображень із кольоровим кодуванням для подальшого аналізу [14]. RedZone об'єднав лазерний і CCTV (Closed-circuit Television) методи для перевірки великих труб. Таким чином, можна отримати більш точну та вичерпну інформацію про стан труб [15]. Системи KARO та PIRAT є іншими методами контролю з використанням кількох датчиків, які можуть автоматично інтерпретувати та класифікувати дефекти в каналізаційних трубопроводах [13, 16]. Проект INNOKANIS представляє новий інструмент шляхом поєднання зум-камери та акустичної технології SewerBatt, яка об'єднує оптичний та акустичний пристрій як альтернативу звичайним відеоспостереженням [16].

Методи спрямовані на виявлення специфічних дефектів відомі завдяки їх здатності надавати кількісні параметри дефектів. Однак дані методи здатні розпізнавати лише 1 чи 2 типи дефектів, але є дорогими у використанні, вимагають висококваліфікованих операторів для інтерпретації даних. Гібридні системи (KARO, PIRAT і SSET) можуть покривати обмеження інших методів, особливо інструментів на основі камер, оскільки використання двох або більше взаємодоповнюючих технологій забезпечує високоякісну та надійну інформацію

про широкий спектр дефектів під час аналізу даних та оцінці каналізаційних трубопроводів. Однак ці технології знаходяться на стадії прототипу і мають високі початкові експлуатаційні витрати. Технології, засновані на використанні звичайної відеозйомки, є найпоширенішими інструментами для інспекції каналізації і можуть надати докази більшості дефектів. Ці методи продуктивні та економічно ефективні, а надану інформацію можна легко аналізувати. Основним їх недоліком є недостатньо якісна інформація про стан труби під водою, про стан підкладки, про інтенсивність дефекту і глибину тріщин, про прогин труби. Крім цього на ефективність аналізу отриманих відеоданих впливає якість зображень і освітлення. Технології сканування і профілювання труби можуть доповнити відеодані і забезпечити якісну інформацію про геометрію і розташування дефектів, однак це додатково ускладнює інтерпретацію і висуває додаткові умови щодо проведення робіт. У табл. 1.1 представлено порівняння вищезгаданих технологій.

Таблиця 1.1 – Порівняння технологій інспекції труб водовідведення

Назва характеристики	На основі відеокамер	Дослідження структури та підкладок (георадар)	Гібридні технології (сканування та оцінювання)	Методи аналізу специфічних дефектів (Електросканування)
Кількість типів дефектів, що виявляються	6	3	8	3
Складність	Низька	Середня	Висока	Середня
Вартість	Середня	Середня	Висока	Середня
Час простою	Високий	Високий	Високий	Низький
Автоматизація аналізу даних	Низька	Середня	Висока	Низька
Потреба в експлуатаційному обладанні	Низька	Середня	Висока	Середня
Якість даних	Середня	Висока	Висока	Висока

Аналіз табл. 1.1 показує, що відеоінспекція забезпечує найкраще співвідношення інформативності і вартості іспекції, а її основним недоліком є порівняно невисока якість даних та рівень автоматизації аналізу даних. Проте ці характеристики можна покращити за рахунок використання сучасних інформаційних технологій інтелектуального аналізу даних, що моделюють когнітивні процеси, притаманні людині під час прийняття рішень. Однак покищо, аналіз відеоданих здійснює сертифікований оператор, кодуючи дефекти і їх характеристики у відповідності до одного зі стандартів. Найбільшої популярності набули Британський стандарт МССS5 [17] та американські стандарт РАСР6 і РАСР7 [18]. Останнім часом з'явилося ряд стартап-компаній, спрямованих на підвищення ефективності аналізу даних відеоінспекції труб водовідведення на основі технологій машинного зору та штучного інтелекту, наприклад SewerAI, Molfar.AI та WinCan [19, 20, 21]. Однак досі аналіз даних не відбувається повністю автоматично і в автономному режимі, результуючі звіти змушені перевіряти кваліфіковані оператори.

Таким чином, основним джерелом інформації є дані відеоінспекції, оскільки їх збір має найменше накладних витрат, а інформативність отриманих даних забезпечує виявлення більшості значимих типів дефектів і визначення їх параметрів. При цьому сучасні засоби аналізу зібраних даних на основі інтелектуальних технологій забезпечують недостатньо високий рівень автоматизації внаслідок невисокої достовірності сформованих звітів. Тому розвиток нових технологій штучного інтелекту, що забезпечують прийнятну точність для практичного використання є актуальним напрямком досліджень.

1.2 Методи, моделі та алгоритми екстракції ознакового опису відеозображень

Ефективність розпізнавання складних візуальних об'єктів залежить від способу ознакового опису зображень [22, 23, 24]. У праці [25] було запропоновано метод виявлення зон інтересу, зокрема тих що містять тріщини, під час інспекції

підземної труби роботом. Запропоноване рішення основане на використанні алгоритму Canny для детектування країв (edge detection), що є класичним методом машинного зору. Даний метод потребує ручного підбору параметрів для конкретного типу труб і не враховує контекстної інформації, що ускладнює його практичне використання. При цьому не було розглянуто питання детального аналізу виявлених зон інтересу для їх кодування в рамках існуючих стандартів інспекції труб.

У працях [26, 27, 28] запропоновано використовувати фільтри Габора для екстракції ознак, що описують спостереження під час інспекції труб водовідведення у поєднанні з класичними моделями машинного навчання (рис. 1.9).

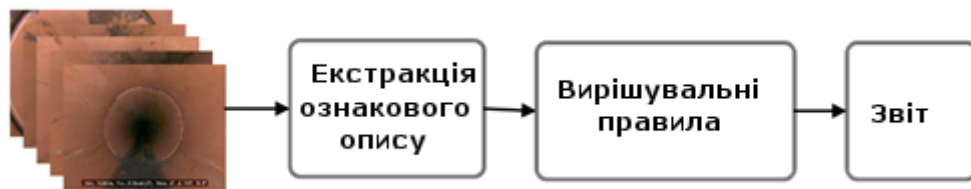


Рисунок 1.9 – Підхід до аналізу даних на основі класичного машинного навчання

Підвищення інформативності екстрактора ознак на основі фільтрів Габора вимагатиме набагато більше ресурсів порівняно з моделями ієрархічної екстракції ознак. При цьому класична модель у вигляді алгоритму випадкового лісу не дозволяє ефективно здійснювати детальний аналіз з одночасним врахуванням контексту. Крім того чітке розділення моделі на екстрактор та класифікатор не дозволяє автоматизувати процес навчання всієї моделі за схемою з кінця-в-кінець.

У працях [29, 30, 31] запропоновано використовувати глибокі моделі аналізу візуальних даних, що забезпечують ієрархічне ознакове подання з високим рівнем інваріантності і інформативності (рис. 1.10)



Рисунок 1.10 – Підхід до аналізу даних на основі глибокого машинного навчання

Ефективність глибокого машинного навчання залежить як від особливостей даних, так і від ефективності вибору архітектури моделі аналізу даних з урахуванням всіх обмежень задачі. Дослідниками в галузі аналізу даних сформульовано ряд універсальних принципів щодо формування ознакового опису даних [32, 33]:

- ієрархічна організація пояснювальних факторів;
- множинність пояснювальних факторів;
- просторово-часова зв'язаність;
- навчання з частковим залученням вчителя;
- спільність факторів під час розв'язання різних задач;
- гіпотеза багатовидів;
- розрідженість ознакового подання;
- гіпотеза про природну кластеризацію;
- простота залежності високорівневого подання факторів.

Поняття, використовувані для опису спостережень, можуть бути визначені в термінах інших більш абстрактних понять, тобто ієрархічно. Тобто високорівневі ознаки формуються методом композиції низькорівневих ознак. Ієрархічні, тобто багатошарові, моделі порівняно з моделями, що мають один прихований шар, мають більшу ємність. Тобто ієрархічні моделі потенційно можуть бути навчені значно більшим абстракціям ознак на верхніх рівнях, кодуючи більш складні залежності в даних при меншій кількості параметрів [34]. За допомогою ієрархічного подання можна ефективно здійснювати або екстракцію інваріантних ознак (*invariant features*), або розділення пояснювальних факторів різного рівня

абстрактності. Розділення пояснювальних факторів може бути корисною під час навчання без вчителя, коли сформований ознаковий опис може повторно використовуватися для вирішення різних завдань.

Вхідні дані описують результат взаємодії множини пояснювальних факторів, що обґрунтовує використання розподіленого подання даних, де кожен параметр може бути повторно використаним для кодування різних вхідних спостережень чи частин вхідного спостереження. У розподіленому поданні даних, на відміну від локальних узагальнень, вхідним сигналом може бути активована більша кількість прихованих змінних. Отже, розподілене подання даних може кодувати більшу кількість різноманітних вхідних конфігурацій даних ніж подання з використанням локального узагальнення.

Під час розробки структури моделі аналізу даних необхідно враховувати обмеження на зміну «крізь час і простір». Спостереження, що сформовані в сусідніх областях простору чи одержані послідовно в часі, повинні бути асоційованими з однаковими значеннями відповідної категорії понять, чи приводити до невеликого зміщення на поверхні багатовиду [35]. Дана властивість може обґрунтовувати використання аугментації даних на основі невеликих випадкових збурень в існуючих навчальних даних, введення обмежень на рецептивні поля моделі відповідно до знань про топологію даних або введення регуляризуючої складової до критерію оптимізації.

Для ефективного використання нерозмічених даних варто враховувати існування статистичного зв'язку між навчанням без учителя та навчанням з учителем. Нерозмічених даних, як правило, багато, тому їх отримання не пов'язано зі значними витратами. Було доведено, що ознакове подання даних \tilde{X} , зручне для обчислення ймовірнісного розподілу $P(\tilde{X})$, є зручним і для навчання $P(Y|\tilde{X})$, де Y – цільова змінна вирішувальних правил [36]. Часткове залучення учителя, що часто відбувається після попереднього навчання без учителя, потрібне для уточнення меж категорій, беручи до уваги, що роздільна гіперповерхня, як правило, проходить через області простору з низькою щільністю ймовірності.

У багатозадачних моделях аналізу даних може досягатися вищий рівень узагальнення внаслідок підсилення статистичного взаємозв'язку між задачами, у яких використовуються спільні пояснювальні фактори [37]. У багатозадачній моделі кожна задача може розглядатися як регуляризуюча складова для інших задач, що змушує алгоритм навчання шукати таку конфігурацію ознакового подання, що задовольняє всі задачі. Тому екстрактор ознак, навчений в рамках багатозадачної моделі, з меншими затратами ресурсів може бути адаптований до вирішення нових чи додаткових задач.

Можливість зниження розмірності ґрунтується на гіпотезі багатовидів, згідно з якою основна щільність ймовірності даних зосереджена біля регіонів, що мають набагато меншу розмірність, ніж оригінальний простір вхідних даних. Крім того ця гіпотеза була узагальнена професором Н. Тішбі, який формалізував поняття про оптимальність ознакового подання у принципі інформаційного пляшкового горла [38]. Згідно даного принципу необхідно знайти баланс між компресією даних та збереженням максимальної предиктивності щодо цільової змінної. Іншими словами реалізація принципу інформаційного пляшкового горла полягає у пошуку компактного ознакового подання \tilde{X} при збереженні достатньої інформації про цільову змінну Y , мінімізуючи наступну цільову функцію

$$J = I(\tilde{X}; X) - \beta I(\tilde{X}; Y), \quad (1.2.1)$$

де $I(\bullet)$ – взаємна інформація між двома змінними;

X, \tilde{X} – дані та їх стиснене ознакове подання;

$\beta > 0$ – параметр компромісу між складністю ознакового подання та кількістю збереженої актуальної для задачі інформації.

Для компресії ознакового подання можна як зменшувати його розмірність, так і стискати множину можливих значень. Наприклад, у працях [39] доводяться регуляризуючі та мета-регуляризуючі властивості дискретного ознакового подання спостережень. При цьому одні дослідники дискретизацію реалізують на певному кроці навчання, а інші здійснюють пряме навчання для формування дискретного ознакового подання [39, 40]. Якщо задати обмеження на вираз (1.2.1) у вигляді

$I(\tilde{X}; X) \leq \gamma$, то принцип інформаційного пляшкового горла зводиться до максимізації взаємної інформації між ознаковим поданням та вихідними цільовими мітками даних $\max_{\tilde{X}} I(\tilde{X}; Y)$.

Для будь-якого даного спостереження x лише мала частина з усіх можливих факторів є значимою, тобто більшість ознак спостереження повинні бути нечутливими до малих змін спостереження x . Тобто більшість екстрагованих ознак повинна бути нульовою. Цю властивість називають розрідженістю (sparsity). Вона може бути досягнута за рахунок різноманітних технік, що ґрунтуються на ефекті редукції причини (explaining away) [41]. Редукція причини полягає у зв'язуванні двох апріорно не зв'язаних причин події, якщо з'являється спостереження цієї події. Властивість розрідженості часто досягають за рахунок використання спеціальних нелінійностей, регуляризації або застосування методів розрідженої компресії в рамках інформаційного пляшкового горла [40].

Під час синтезу моделей аналізу даних багато дослідників використовують гіпотезу природньої кластеризації. Для ознакового подання даних \tilde{X} щільність ймовірності $P(\tilde{X} | Y = k)$ для різних категорій k прагне розділятися, а не накладатися. Іншими словами категорії асоціюється з різними багатовидами. Локальні зміни в багатовидах мають тенденцію до збереження категорії.

В ефективних ознакових поданнях даних фактори зв'язані один з одним через прості лінійні або дистанційні залежності. Дане твердження обґрунтовує вибір гіперпараметрів екстрактора ознакового опису за результатами оцінки достовірності вирішувальних правил.

У працях [42] було запропоновано використовувати кодування міток завадозахищеними кодами з матриці Адамара, надаючи ознаковому поданню властивостей кодів, що виправляють помилки. Було показано, що в результаті вдалося підвищити узагальнюючі властивості глибокої нейронної мережі. Рішення щодо віднесення вхідного зразка до однієї з категорій здійснювалося за відстанню Хеммінга. Одним з недоліків даного підходу можна вважати, що він не враховує

внутрішньої структури даних, залежностей між класами і нав'язує невластиву для них структуру.

Найбільший успіх в задачах інтелектуального аналізу зображень було отримано в рамках технології згорткових нейронних мереж. В даній мережі згідно принципу просторової зв'язаності кожен нейрон має обмежене рецептивне поле і послідовно сканує зображення, а не сприймає його повністю, тобто не зв'язаний з всіма пікселями. Це дозволяє значно зменшити кількість параметрів і обчислювальну складність моделі екстракції ознакового опису за рахунок повторного використання вагових коефіцієнтів нейронів. З 1989 року по сьогодні було виконано ряд архітектурних вдосконалень згорткової нейронної мережі. Вдосконалення стосувалися таких аспектів як оптимізація параметрів, регуляризація, структурне переформулювання тощо. Проте останні тенденції в модернізації згорткової мережі пов'язані з реструктуризацією існуючих компонентів обробки даних та проектуванням нових блоків і модулів (рис. 1.11). Залежно від типу архітектурних модифікацій згорткові мережі можна класифікувати на сім різних груп [43]:

- експлуатація простору (spatial exploitation), що основана на використанні різномасштабних фільтрів;
- експлуатація глибини (depth exploitation), тобто збільшення кількості шарів і міри щодо забезпечення незатухання градієнту на нижніх шарах;
- експлуатація мульти-шляхів (multi-path exploitation), тобто використання міжшарових з'єднань та псевдо-ансамблювання на основі dropout і drop-connection;
- експлуатація ширини (width exploitation), тобто міри щодо збільшення ширини нейронної мережі за рахунок мульти-з'єднань;
- експлуатація карти ознак (feature map exploitation), тобто підвищення інформативності ознак за рахунок їх селекції/зважування на основі Squeeze-Excitation-модуля;
- експлуатація підсилення каналів (channel boosting exploitation), тобто додавання штучних каналів до оригінальних каналів, що сформовані за рахунок

переносу знань (transfer learning), використання генеративних моделей навчання або агрегації частинних описів спостережень (на рівні модельностей);

– використання модулів уваги (attention based CNNs), тобто використання м'яких масок для карти ознак з метою фокусування моделі на важливих регіонах зображення/карти ознак.

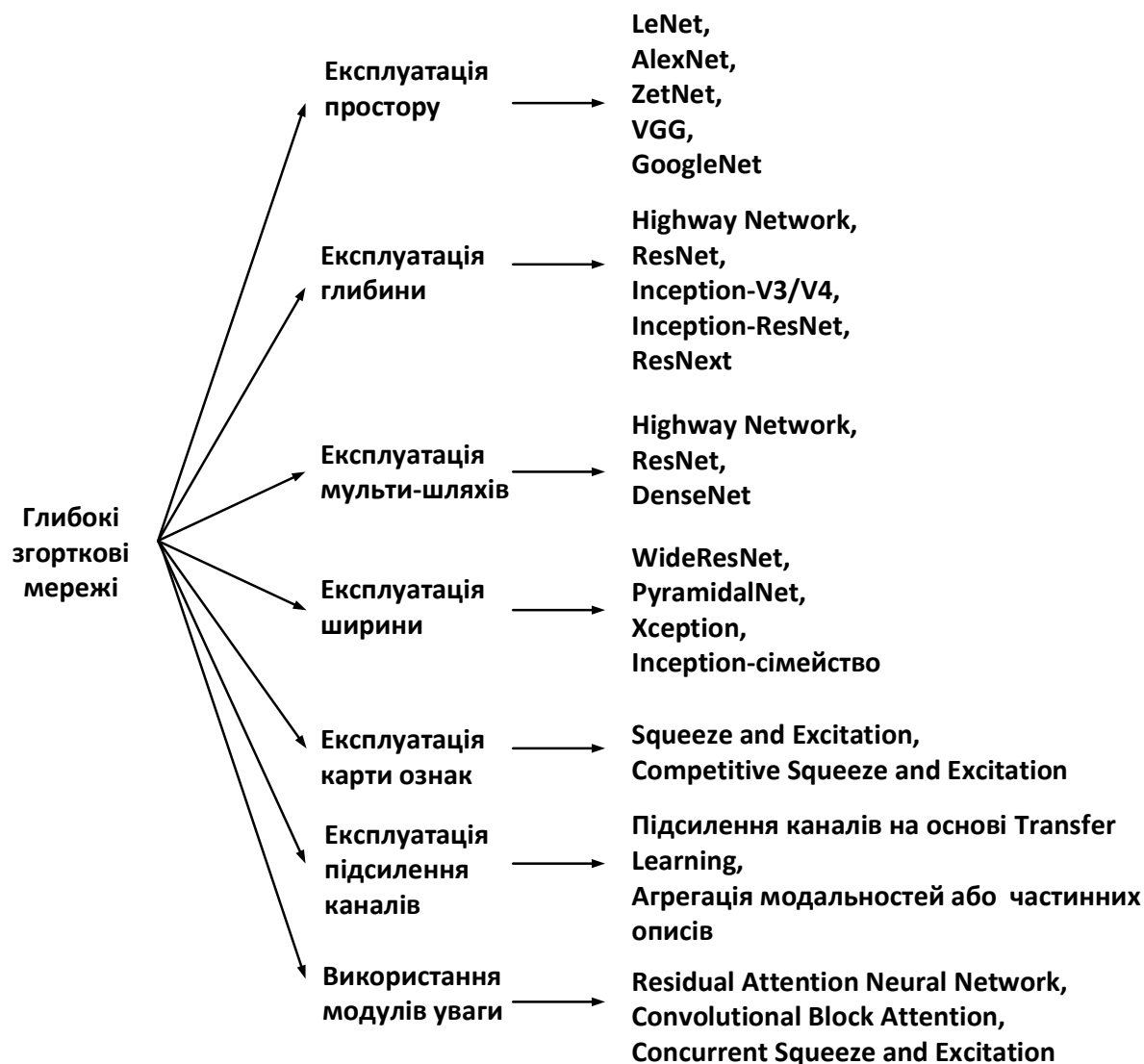


Рисунок 1.11 – Таксономія багатошарових згорткових нейронних мереж за способом їх удосконалення

Для підвищення ефективності нейронних мереж часто використовується механізм уваги (Attention) та внутрішньої уваги (self-attention) [44], що дозволили враховувати взаємозв'язки між віддаленими частинами вхідних даних та

проміжними результатами обробки цих даних. Подальший розвиток цих механізмів привів до появи нових будівельних блоків – так званих трансформерів. Трансформер – це поставлені один за одним модулі уваги, які дають змогу вхідну послідовність векторів трансформувати в нову послідовність векторів, які кодують інформацію про контекст кожного елемента [45]. При цьому вхідні елементи надходять у вигляді векторного подання (embedding), до якого додається вектор позиційного кодування (positional encoding) для збереження інформації про порядок елементів. Для фокусування на різних аспектах вхідної інформації використовується декілька паралельних модулів внутрішньої уваги (Multi-Head Attention), що формують множину підпросторів представлень (representation subspaces).

Вхідні елементи трансформера називають токенами, які у задачах машинного зору є патчами зображення або патчами високорівневих карт ознак, що витягуються у вектори перед подачею на модуль внутрішньої уваги. У традиційних трансформерах розмірність токена на вході і виході не змінюється, але у машинному зорі існує необхідність в аналізі більш дрібного контексту в нижчих шарах і більш глобального контексту у верхніх шарах моделі. Тому у Swin-трансформерах, перший енкодер здійснює аналіз маленьких патчів 4x4 пікселі, а перед подачею на наступні енкодери відбувається об'єднання менших патчів в більші (Patch Merging) [46]. Оскільки найбільш обчислювально складним елементом трансформаторів є модуль уваги, що не розпаралелюється, в Swin-трансформерах увага для кожного токена обчислюється не до всіх сусідів, а лише до сусідів із фіксованого вікна. Для запобігання втрат репрезентативної здатності після кожного блоку з віконними модулями уваги (Window Multi-Head Attention) використовується аналогічний шар зі зміщеними по діагоналі вікнами уваги.

Трансформери завдяки здатності моделювати довгострокові залежності забезпечують ефективну екстракцію глобальних ознак та меншу індуктивну упередженість. Тому трансформерам вдається досягти порівняної, або навіть більшої точності на великих наборах даних. Однак навіть сучасні вдосконалення візуальних трансформерів, такі як ViT, PiT та SWIN все ще поступаються

згортковим мережам за обчислюваною складністю і за узагальнюючою здатністю на вибірках обмеженого обсягу [46, 47].

У працях [29, 30, 48] показано результати застосування згорткових мереж для детектування дефектів труб водовідведення. Однак отримані результати носять модельний характер, оскільки використовується обмежений алфавіт класів, що не може покрити жоден із популярних стандартів формування звіту про інспекцію труб водовідведення і ігнорується просторово-часовий контекст. Шум, тимчасові втрати видимості, схожість різних дефектів між собою залежно від орієнтації камери обумовлюють необхідність моделювання міжкадрових залежностей, що відноситься до задач аналізу послідовностей. В задачах аналізу послідовностей (тобто процесів, що розгорнуті в часі) лідерами є рекурентні та темпоральні згорткові нейронні мережі.

Серед рекурентних мереж найбільш популярними є LSTM та GRU-мережі, які складаються з комірок пам'яті [49]. Ці комірки з'єднані як послідовні шари, що дозволяє мережам вибірково запам'ятовувати або забувати вхідні навчальні дані через унікальну структуру з вентилями. Під час навчання мережі оптимізуються параметри цих вентилів. Рекурентні мережі мають ряд суттєвих проблем, до яких відносять чутливість до перших зразків даних на вході, складний характер динаміки навчання з можливим переходом до хаотичної поведінки і складність розпаралелювання внаслідок послідовної організації функціонування моделі.

Темпоральні згорткові мережі використовують локальні рецептивні поля для сканування сигналів з 1D-топологіями. Згортки темпоральних мережу можуть бути регулярними або казуальними [50]. Регулярні згортки є симетричними, в той час як в казуальній темпоральній згортковій мережі рецептивне поле вузлів охоплює поточні і минулі дані (рис. 1.12). Тобто результат аналізу кожного вузла залежить тільки від значень сигналу у минулих відліках часу. При цьому для збільшення рецептивного поля моделі її необхідно стекувати в глибину.

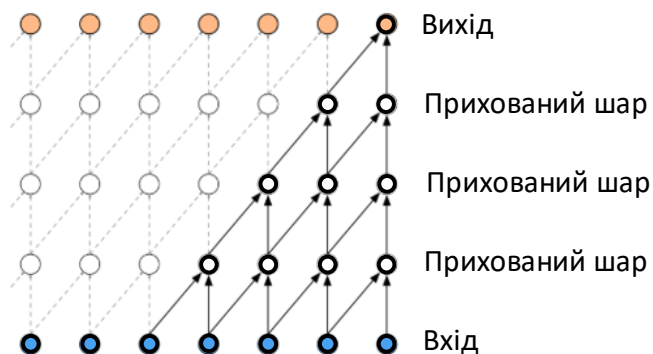


Рисунок 1.12 – Архітектура каузальної моделі для аналізу послідовностей

Для аналізу даних високої розмірності з відомою топологією можна використовувати діряві (dilated) рецептивні поля, які застосовуються до локальних ділянок даних, розмір яких перевищує кількість входів кодера. Тобто частина даних, що накривається рецептивним полем, ігнорується, а аналізу піддаються дані з певним просторовим або часовим кроком (рис. 1.13).

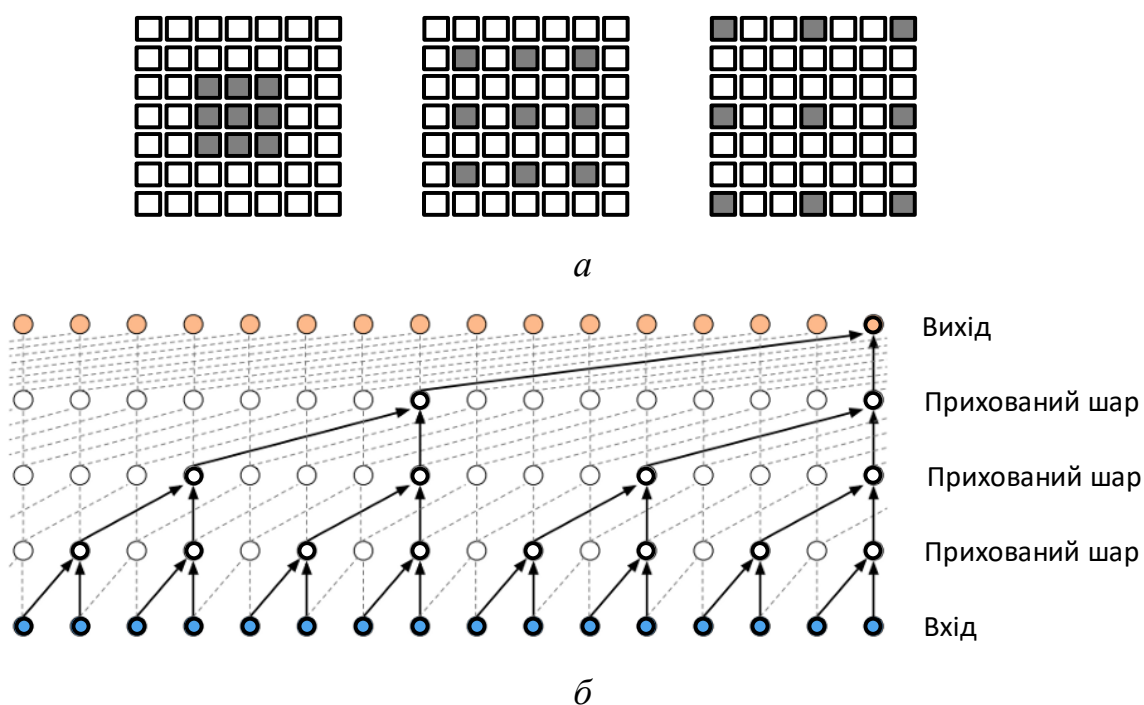


Рисунок 1.13 – Діряві рецептивні поля моделі аналізу даних :

a – 2D-топология; b – 1D-топология каузальної моделі

Стекування вузлів з дірливими рецептивними полями дозволяє досягти великого рецептивного поля моделі з використанням невеликої кількості шарів і параметрів. При цьому не відбувається втрат інформації, оскільки кодер здійснює сканування вхідних даних в просторі або часі. Крім того наявність пропусків дозволяє знизити чутливість до високочастотних складових шуму і, як результат, підвищити завадозахищеність моделі.

Останні дослідження показали, що аналіз послідовностей згортковими мережами в рамках 1D-топології з модифікованими рецептивними полями, які називають дірливими (dilated), перевершують рекурентні мережі як в опертивності, так і в ємності [51]. А застосування механізму уваги (attention) у вигляді додаткового модуля, що обчислює маску для фокусування на важливих просторових та канальних ознаках і придушення неважливих ознак, дозволяє додатково підвищити потужність ознакового подання. При цьому згорткові моделі подання даних є зручними для застосування різноманітних технік і протоколів навчання.

Відомо, що чим менший обсяг розмічених даних, тим помітніше ефективність навчання залежить від початкової ініціалізації вагових коефіцієнтів. Саме це і спонукало розвиток методів переносу знань (Transfer Learning). Ще одним підходом до підвищення ефективності початкової ініціалізації вагових коефіцієнтів мереж є використання різних методів самонавчання (unsupervised learning, self-supervised learning). Методи самонавчання можна поділити на три основні типи : генеративні, контрастні та генеративно-контрастні.

Генеративні методи самонавчання здійснюють навчання кодера (екстрактора ознак) кодувати вхідний зразок x у вектор прихованих змінних z та декодера для реконструювання x з z . Характерними представниками генеративних методів є знешуюмлюючий автоенкодер (Denoising Autoencoder) та варіаційний автоенкодер (Variational Autoencoder) [52]. Здатність відновлювати оригінальний розподіл даних без припущень про наступні задачі забезпечує ефективність генеративних моделей для широкого кола задач. Однак генеративні моделі досить чутливі до

рідкісних зразків даних і моделюють розподіл на рівні пікселів, а не на рівні абстракцій, що більш характерно для задач класифікаційного аналізу даних.

Метою контрастного навчання є навчання порівнянню на основі шумового контрастного оцінювання, що виражено у цільовій функції моделі

$$J = E_{x, x^+, x^-} \left[-\log \left(\frac{e^{f(x)^T f(x^+)}}{e^{f(x)^T f(x^+)} + e^{f(x)^T f(x^-)}} \right) \right], \quad (1.2.2)$$

де x^+ – зразок даних, що подібний до x ;

x^- – зразок даних не подібний до x ;

$f(\bullet)$ – функція, що описує екстрактор ознак.

Для більшої кількості несхожих пар можна оцінювати взаємну інформацію між позитивною парою і асоційованими з нею K негативних пар, так звану InfoNCE [53],

$$J = E_{x, x^+, x^k} \left[-\log \left(\frac{e^{f(x)^T f(x^+)}}{e^{f(x)^T f(x^+)} + \sum_{k=1}^K e^{f(x)^T f(x^k)}} \right) \right]. \quad (1.2.3)$$

Сучасні фреймворки контрастного самонавчання поділяються на два типи : контекст-зразок контраст (context-instance contrast) та зразок-зразок контраст (instance-instance contrast). Обидва типи значно перевершують ефективність навчання екстрактора ознак порівняно з генеративними моделями для задачах класифікаційного аналізу, оскільки мета контрастного навчання природньо відповідає задачам класифікаційного аналізу даних. Крім того, контрастне навчання обчислювально ефективніше, оскільки відповідає необхідність у використанні декодера. Дослідниками було розроблено такі відомі методи контрастного самонавчання як MoCo, SimCLR, BYOL та SwAV [54], які продемонстрували високі результати на різних бенчмарках машинного зору. Проте контрастне навчання використовується для формування ознакового опису в неперервному просторі, а питання формування дискретного ознакового подання досі є малодослідженим. Крім того неналежне налаштування гіперпараметрів часто призводить до деградації результатів і формування однакового ознакового опису

для всіх зразків, особливо у випадку модифікації для усунення негативних зразків з процесу навчання.

Таким чином, в існуючих дослідженнях щодо синтезу глибоких моделей для аналізу даних відеоінспекції труб водовідведення ігнорується інформація про просторово-часовий контекст і кожен кадр аналізується незалежно. Тому розробка моделей, що моделюють міжкадрові залежності, для відновлення неповної інформації на певних відеокадрах є актуальною задачею. При цьому найбільш перспективним підходом до синтезу екстрактора ознакового опису є використання згорткової нейронної мережі, що ініціалізується шляхом констратного самонавчання.

1.3 Методи, моделі та алгоритми ідентифікації об'єктів інтересу на зображеннях відеоінспекції

Для коректного кодування функціонального стану труб водовідведення згідно існуючих стандартів необхідно мати як інформацію про тип дефектів, так і про їх локалізацію, орієнтацію, та близькість до вхідних-вихідних терміналів (колодців) і відгалужень. Тому у працях [28, 29] запропоновано використовувати детектори об'єктів інтересу, що виділяють дефекти обмежувальними рамками. В задачах детектування об'єктів інтересу на зображеннях як правило використовують глибокі нейронні мережі з підмережею детектора, яка повертає дані про локалізацію об'єктів та впевненість розпізнавання для кожного класів. При цьому сучасні моделі детектування об'єктів на основі згорткових нейронних мереж розвиваються паралельно в двох різних напрямках [55] :

- модель з генератором пропозицій областей інтересу з наступною їх класифікацією на категорії об'єктів;
- модель, що вирішує задачу детектування шляхом прямого вирішення задач класифікації та регресії для прогнозування категорій і обмежувальних прямокутників об'єктів інтересу відповідно.

У працях [56] розглядаються моделі детекторів з генератором пропозицій регіонів, до яких відносяться R-CNN, SPP-мережа, Fast R-CNN, Faster R-CNN, R-FCN, FPN та Mask R-CNN. В перших моделях в рамках даного підходу пропозиції регіонів формувалися алгоритмом Selective Search [57], що генерував 2000 різноманітних ділянок, які з найбільшою ймовірністю містять об'єкти інтересу. Для підвищення швидкодії у нових версіях алгоритм Selective Search було замінено на мережу вибору ділянок-кандидатів (RPN – region proposal network), в результаті чого було створено архітектуру Faster R-CNN. На вхід даної мережі подається ділянка розміром $n \times n$, що взята з повнорозмірної карти ознак, результат передається на два повнозв'язні шари: регресор обмежувальної рамки (box-regression) та класифікатор обмежувальної рамки (box-classification). Ділянки-кандидати, отримані за допомогою RPN, представлені координатами обмежувального прямокутника і ймовірністю знаходження об'єкта в даній області, що обчислена із застосуванням функції softmax. Крім того набула поширеного використання мережа піраміди ознак (Feature Pyramid Net, FPN). FPN забезпечує поєднання переваг високої роздільної здатності ознакових карт нижніх шарів екстрактора ознак і переваг високої узагальнюючої і семантичної спроможності ознакових карт верхніх шарів [58].

До методів, оснований на регресії і класифікації, відносяться MultiBox, AttentionNet, G-CNN, SSD (Single Shot Detector), YOLO (You Only Look Once), DSSD (Deconvolutional Single Shot Detector) та DSOD (Deeply Supervised Object Detector) [59]. В моделях з архітектурою SSD вхідне зображення проходить через ряд згорткових шарів, в результаті чого формується набір карт ознак різного масштабу. Для кожної точки кожної карти ознак застосовується згортковий фільтр розміром 3×3 для отримання множини обмежувальних прямокутників. Для кожного обмежувального прямокутника одночасно оцінюються просторові зміщення і ймовірність знаходження об'єкта. У процесі навчання реальні обмежувальні прямокутники співставляються з прогнозованими для виключення хибних спрацювань. З метою зменшення кількості пропонованих обмежувальних прямокутників різного масштабу використовується придушення немаксимумів, що

забезпечує об'єднання схожих один на одного прямокутників в один. Крім того використовується метод відбору найскладніших пар (hard negative mining) [60], відповідно до якого на кожній ітерації навчання використовується лише частина негативних зразків.

Дещо інший підхід до побудови детектора реалізовано в архітектурі Yolo, яка описана в працях [61]. В алгоритмі YOLO зображення поділяється сіткою на комірки, для кожної з яких оцінюється ймовірність присутності об'єкта взагалі. Потім будуються кілька найімовірніших положень об'єкта у вигляді прямокутників з центром у комірці. Для кожного отриманого прямокутника виконується оцінювання ймовірностей наявності у ньому об'єктів кожного аналізованого класу. Оцінка ймовірності знаходження об'єкта конкретного класу в поточному обмежувальному прямокутнику – це добуток оцінки ймовірності знаходження об'єкта в комірці і оцінки ймовірності для конкретного класу. У випадку YOLOv3 для екстракції ознак використовується згортова нейронна мережа, яка складається з 53 шарів, де як фільтри використовуються згортки розміром 3×3 і 1×1 та Residual-блоки, які додають до виходу поточного шару значення з виходу попереднього шару.

Існує ціла низка модифікацій і удосконалень базових архітектур Faster R-CNN, Yolo та SSD, які постійно удосконалюються і розвиваються. Однак для ефективного навчання даних моделей необхідно мати великий обсяг розмічених навчальних даних у вигляді маркованих обмежувальних рамок. При цьому моделі детекторів чутливі до цілого ряду незбалансованостей : незбалансованість класів, незбалансованість масштабів, просторова незбалансованість та незбалансованість функції втрат [62]. Існують різні методи зниження чутливості до різних видів незбалансованості, проте основним джерелом незбалансованості є дані. Ситуація ще більш ускладнюється, якщо екстрактор ознак ініціалізований шляхом переносу знань (Transfer Learning) від задачі, що не має статистичного зв'язку з вирішуваною задачею.

У працях [63] пропонуються методи візуальної одометрії для визначення позиції та орієнтації камери відеоінспекції в системі водовідведення з метою

підвищення рівня автономності збору інформації, а також з метою врахування орієнтації камери під час визначення координат дефектів. На практиці ігнорування орієнтації камери може дезорієнтувати стосовно локалізації дефектів, оскільки камеру часто повертають для детальнішого огляду зон інтересу або для уникнення занурення під воду чи для проходження через завади.

Зміна орієнтації камери збільшує невизначеність координат та внутрішньокласову варіативність візуальних спостережень дефектів, а також перетин класів різнотипних дефектів в просторі ознак. Традиційним підходом до зниження варіативності і відповідно ентропії даних є використання деревовидної структури. У працях [64, 65] доведено ефективність використання ієрархічної композиції моделей глибокого навчання для класифікаційного аналізу текстової інформації. У праці [67] запропоновано використовувати ієрархічну композицію моделей в медичній діагностиці, детекторах шкідливого програмного забезпечення та класифікаторі дефектів труб водовідведення для відокремлення спостережень нормального функціонального стану від інших. Введення ієрархічності окрім зменшення варіативності даних на нижніх рівнях моделі може забезпечувати збалансованість навчальних даних та зменшувати обчислювальну складність в режимі екзамену.

У праці [28] для аналізу зображення відеоінспекції здійснювався його поділ сіткою на комірки, координати яких наперед відомо (рис. 1.14). Передбачалося, що при фіксованій орієнтації камери комірки зображення можна співставити з координатами на трубі. При цьому не всі комірки однаково придатні для аналізу, оскільки значення одометра актуальне лише для бічних комірок. Основним недоліком підходу щодо роздільного аналізу комірок є ігнорування контекстної інформації та припущенні про наперед відому орієнтацію камери.



Рисунок 1.14 – Розбиття зображення відеоінспекції на патчі для роздільного аналізу

У працях [28, 29] зазначено, що під час синтезу класифікатора функціонального стану труб водовідведення необхідно враховувати наявність шумової складової. Дані відео інспекції труб містять значну кількість артефактів та шумів, наприклад, засвічування поверхні, бризки води, тимчасові неполадки відеопристрою, дефекти зображень, схожість бруду та піни на реальні дефекти. У працях [63] доведено, що підвищення стійкості моделей аналізу даних до шуму та змагальних атак можливо за рахунок заміни softmax шару на дистанційні функції належності та вирішувальні правила. Отримані результати перевершують результат відомих методів змагального навчання (Adversarial training). Для тренування таких моделей на розміченому навчальному наборі можна використовувати контрастну $J_{contrast}$ або триплетну $J_{triplet}$ функції втрат, або їх модифікації [64] :

$$J_{contrast} = [(1 - y)d(z_a, z_b) - yd(z_a, z_b) - \alpha]_+ \quad (1.3.1)$$

$$J_{triplet} = [d(z_a, z_n) - d(z_a, z_p) + \alpha]_+ \quad (1.3.2)$$

де y – індикатор, що вказує чи належать зразок x_a та x_b до одного і того ж класу;

z_a, z_b, z_n, z_p – ознакове подання зразків x_a, x_b, x_n, x_p на виході екстрактора ознак;

$$[t]_+ = \max(t, 0);$$

$d(\bullet)$ – дистанційна міра, наприклад евклідова або косинусна;

α – константа відступу.

Сіамська мережа складається з двох або трьох екстракторів ознак зі спільними параметрами (рис. 1.15). Параметри моделі W вважають оптимальними, якщо відстань між ознаковим описом на виході мережі для пари семантично схожих зразків є меншою ніж відстань між ознаковим описом семантично відмінних зразків. Тобто навчання сіамських мереж можна вважати навчанням метрики подібності (metric learning). Навчена сіамська мережа може виконувати роль дистанційної метрики для «лінивих» алгоритмів побудови вирішувальних правил, таких як метод k -найближчих сусідів. Також дані мережі можуть розглядатися як адаптивне ядро для ядерних алгоритмів побудови вирішувальних правил.

Одним із традиційних підходів до підвищення робастності класифікаційних алгоритмів є побудова ансамблевих моделей. Однак, як було показано в працях [65, 66], для уникнення ефекту перенавчання ансамблевої моделі слід подбати про диверсність (різноманіття). Під диверсністю, як правило, розуміють кількісну оцінку різниці допущення однакових помилок серед моделей в ансамблі. Тому диверсність інколи оцінюють на основі виразу

$$\bar{D} = E - \bar{E}, \quad (1.3.3)$$

де \bar{D} – усереднена оцінка диверсності моделей ансамблю;

E – загальна помилка рішень, прийнятих за результатами голосування моделей ансамблю;

\bar{E} – середня помилка рішень, прийнятих індивідуальними моделями ансамблю.

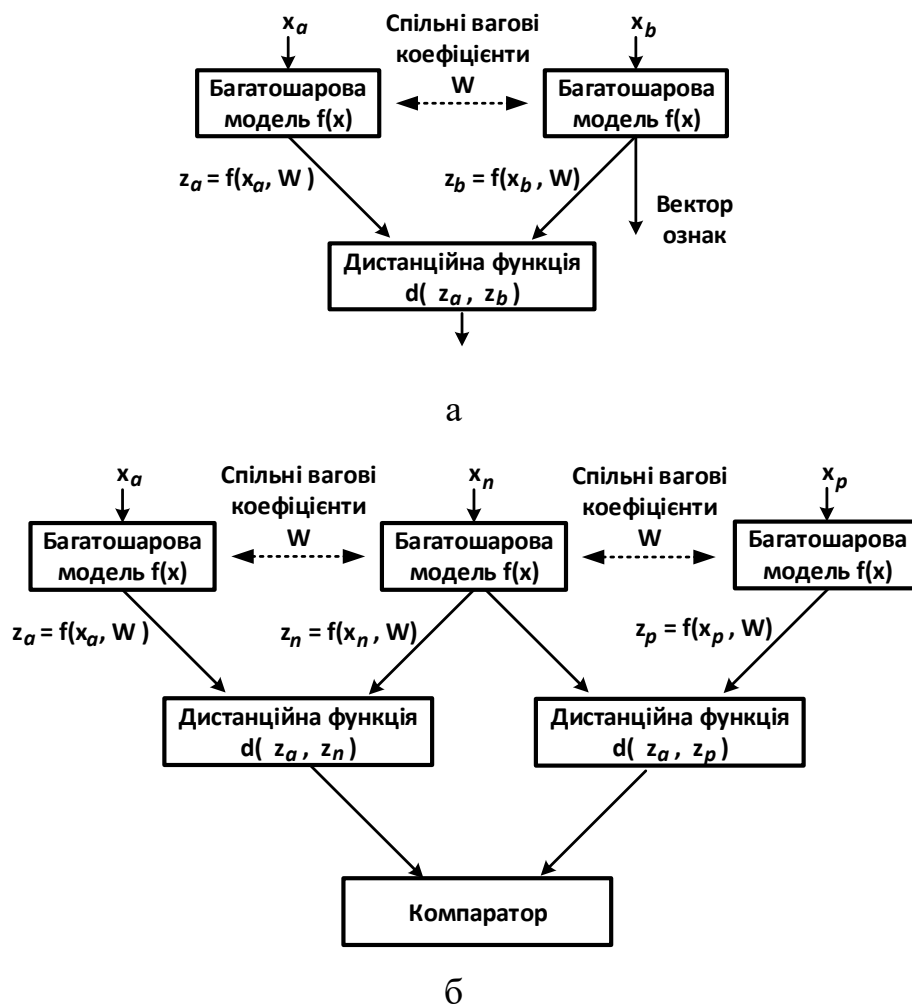


Рисунок 1.15 – Структура сіамської моделі аналізу даних :

а – модель для навчання з функцією втрат (1.3.1) ;

б – модель для навчання з функцією втрат (1.3.2)

Існує багато визначень диверсності та способів оцінювання її рівня. Один з підходів до оцінювання диверсності полягає в аналізі ймовірностей одночасної незгоди моделей ансамблю під час прийняття рішень. В інших підходах диверсність моделей ансамблю розглядають як оцінку структурної різноманітності в ансамблі, що може бути виміряна ентропійними та інформаційними мірами [67].

Одним із методів забезпечення диверсності є використання в ансамблі моделей, що основані на різних принципах. Однак також популярним підходом до забезпечення диверсності є введення елемента рандомізації [66] :

– маніпуляції з навчальними вибірками моделей (формування випадкових підвбірок, ініціалізація вагових коефіцієнтів навчальних зразків);

- вибір різних підпросторів (випадкова вибірка ознак);
- маніпуляції з параметрами навчання (ініціалізація нейронної мережі випадковими числами, введення регуляризуючого штрафу за кореляцію результатів мережі з результатами інших мереж ансамблю);
- маніпуляції з поданням вихідних значень моделей (подання класів кодами, що виправляють помилки, або випадкові зміни біт класів в деяких навчальних зразках).

Основним недоліком ансамблевих алгоритмів є підвищення ресурсомісткості навчання та екзамени моделі та необхідність вживання додаткових мір щодо диверсності. Питання ефективності використання ресурсів є важливим у випадку розгортання моделей в хмарному середовищі, де білінг виконується за принципом pay-as-you-go.

Класичним методом підвищення точності багатокласової класифікації є її заміна на еквівалентну множину двох-класових задач. Одним з найбільш ефективних методів зведення багатокласової класифікації до серії двохкласових є двійкове кодування біт класів кодами, що коригують помилки (Error Correcting Output Codes, ECOC) [68].

Під час використання кодів, що коригують помилки, номер класу записують у вигляді k -значного двійкового числа. Для цього здійснюють навчання k -класифікаторів, кожен із яких розпізнає один із k розрядів номера класу. За результатами розпізнавання вхідного вектора кожним із класифікаторів однозначно відновлюють номер класу, до якого він належить [66]. Якщо окремі класифікатори помиляються, то номер класу відновлюють методом заміни одержаного номера номером класу, що найближчий до одержаного за метрикою Хеммінга (рис. 1.16).

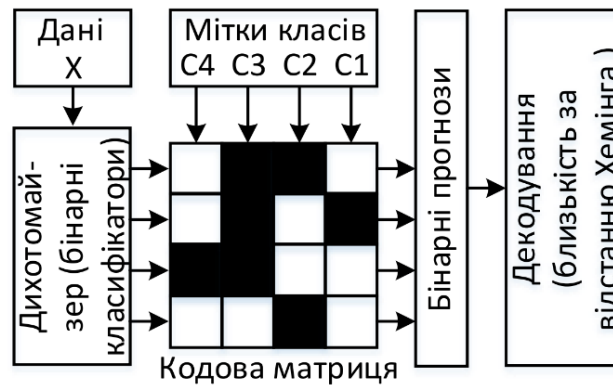


Рисунок 1.16 – Ілюстрація до схеми багатокласової класифікація з використанням двійкового кодування номера класу кодами, що виправляють помилки

На рис. 1.16 показано кодову матрицю, де чорні й білі комірочки позначають мітки 0 або 1 для зразків відповідних класів на етапі навчання бінарних класифікаторів. У режимі екзамену цю матрицю використовують для порівняння прогнозованого коду з кодами класів. Ефективна кодова матриця повинна забезпечувати найбільшу відстань Хеммінга як між рядками кодової матриці, так і між стовпцями. Відстань між рядками забезпечує можливість самовиправлення помилок, а відстань між стовпцями забезпечує некорельованість результатів кожного з бінарних класифікаторів, що навчаються. При цьому існує багато способів формування номера класу для кодової матриці, однак саме коди, що виправляють помилки, (самокорегувальні коди) забезпечують кодування з необхідними властивостями. Поширеним способом формування самокорегувальних кодів є використання матриць і кодів Адамара [6].

Основним недоліком методу ЕСОС є ігнорування структури класів під час побудови кодової матриці без можливості оптимізації коду кожного класу в процесі навчання. Крім того, відсутність кодового радіуса для кожного номера класу, що вказує на кратність помилок, які можуть бути виправлені, ускладнює виявлення викидів або новизни в даних.

Одним із шляхів підвищення точності класифікаторів з одночасним зменшення їх обчислювальної складності є використання методу дистиляції знань

(Knowledge distillation) [47, 66]. Суть методу полягає у попередньому навчанні великої моделі (учителя), який в подальшому може використовуватися для навчання меншої моделі (учня). На виході методі вчителя формується розподіл ймовірності належності до кожного з класів, що є більш інформативною міткою вхідних даних ніж one-hot-код з позначенням 100% належності лише одному класу. На виході моделі учителя формується більше інформації, враховується взаємозв'язок між класами, тому навчання учня відбувається більше ефективно навіть попри меншу кількість параметрів. Недоліком методу є потреба у навчання великої моделі вчителя на великому обсязі даних.

На усунення недоліків відомих методів машинного навчання та розпізнавання образів спрямовано інформаційно-екстремальну інтелектуальну технологію аналізу та синтезу здатних навчатися діагностичних систем, що ґрунтується на максимізації інформаційної спроможності системи розпізнавання шляхом введення в процесі навчання додаткових інформаційних обмежень [69]. Ця технологія ґрунтується на принципах :

- двійкове кодування міток класів розпізнавання двійковими прототипами для реалізації принципів самокоректуючих кодів (Error-Correcting Output Coding);
- трансформація простору ознак для зменшення відстані між зразками однакових класів та збільшення відстані між зразками різних класів в рамках двійкового простору Хеммінга (рис. 1.17);
- оптимізація в інформаційному розумінні вирішувальних правил, що відновлюються в радіальному базисі простору Хеммінга, з метою підвищення стійкості до шуму і новизни в даних за рахунок збільшення буферної зони між класами та рахунок врахування габаритів розподілу кожного класу;
- прийняття рішень у рамках ІЕІ-технології здійснюється в рамках детерміновано-статистичного підходу шляхом побудови відносно простого детермінованого класифікатора, статистична корекція якого здійснюється в процесі навчання з метою побудови безпомилкових, за навчальними матрицями, вирішувальних правил;

– методи ІЕІ-технології базуються на вибіркового підході математичної статистики і орієнтовані на застосування прийнятних з практичних міркувань обсягів репрезентативних навчальних вибірок.

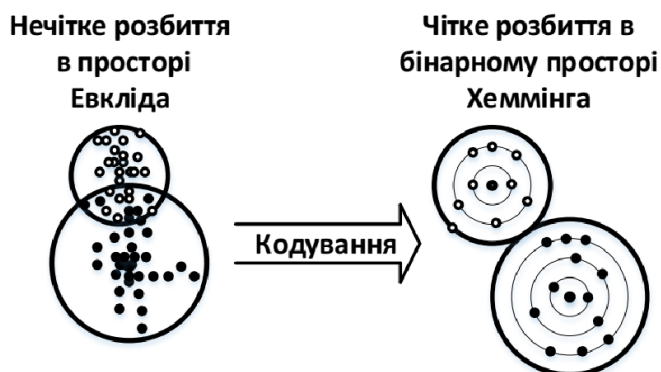


Рисунок 1.17 – Ілюстрація для пояснення ідеї трансформації простору ознак в рамках ІЕІ-технології

В рамках ІЕІ-технології процес навчання моделі класифікаційного аналізу полягає в реалізації ітераційної або багатоетапної процедури оптимізації генотипних та фенотипних параметрів функціонування, що впливають на функціональну ефективність навчання [70]. До генотипних параметрів відносяться параметри екстрактора ознакового опису спостережень, включаючи етап дискретизації, що забезпечує реалізацію принципу інформаційного пляшкового горла. До фенотипних параметрів відносять двійкові прототипи класів та кодові радіуси контейнерів класів. Як інформаційний критерій використовується нормовані модифікації мір К. Шенона або С. Кульбака, що виражені як функція від чутливості, специфічності та частоти помилок першого і другого роду. Екстрактор ознакового опису в рамках ІЕІ-технології може бути побудований в рамках різних нейромережових підходів, однак питання такої інтеграції потребує додаткових досліджень [22, 66].

Таким чином, за наявності апріорної інформації про орієнтацію камери під час інспекції в трубі водовідведення задачу ідентифікації дефектів можна звести до класифікаційного аналізу фіксованих ділянок зображення, що відповідають

актуальним значенням одометра. При цьому орієнтація камери може визначатися на першому рівні ієрархічної моделі. Зашумленість даних інспекції обумовлює актуальність дослідження ефективності використання та поєднання ідей і методів композиції моделей, сіамських мереж, кодів, що виправляють помилки або інформаційно-екстремальної інтелектуальної технології для задач розпізнавання функціонального стану труб водовідведення.

Висновки до першого розділу

Аналітичний огляд сучасного стану та тенденцій розвитку систем моніторингу підземного труб водовідведення дозволяє зробити наступні висновки:

1. Для інспекції труб водовідведення використовують відеокамери, в тому числі тепловізійні, сонари, георадари, лазерні сканери, електросканери та їх комбінації з метою виявлення дефектів та оцінювання цілісності стінок труб та стану ґрунту навколо труб. Проте основним джерелом інформації є відеокамери, оскільки відеоінспекція вимагає найменше накладних витрат, а інформативність отриманих даних забезпечує виявлення більшості важливих типів дефектів. При цьому основна тенденція розвитку сучасних технологій аналізу даних відеоінспекції пов'язана з впровадженням технологій машинного зору та штучного інтелекту. Однак дані технології досі не забезпечують високий рівень автоматизації внаслідок недостатньої достовірності сформованих звітів, тому удосконалення інтелектуальних технологій аналізу даних відеоінспекції є актуальним напрямком досліджень.

2. Сучасні методи формування ознакового опису багатовимірних даних спираються на ряд принципів: розподілене кодування; ієрархічна структура ознакового подання; навчання з неповним залученням вчителя; зниження розмірності даних згідно гіпотези багатовидів; природна кластеризація зразків кожного класу у високорівневих ознакових просторах; часова і просторова зв'язаність спостережень; розрідженість ознакового подання; простота залежності факторів у високорівневому ознаковому поданні. Найбільш популярними

архітектурами екстракторів ознак з відеозображень є згорткова та трансформаторна нейронні мережі, проте згорткова архітектура є обчислювально ефективнішою і менш вибагливою до обсягу навчальних даних. При цьому під час відеоінспекції труб актуальними проблемами формування ознакового опису є врахування просторово-часового контексту і впливу шумових факторів, а також забезпечення ефективності навчання за умов обмеженого і незбалансованого обсягу розмічених даних.

3. Для ідентифікації об'єктів інтересу на зображеннях відеоінспекції, як правило, використовуються різноманітні моделі і методи детектування, які є досить чутливими до різного роду незбалансованості і потребують трудомісткого процесу розмітки навчальних даних прямокутниками зон інтересу. Проте декомпозиція задачі аналізу відеозображень і врахування апріорної інформації про процес інспекції може звести задачу ідентифікації об'єктів інтересу до ряду більш простих задач класифікації. Наприклад, введення моделі контексту може зменшити варіативність і незбалансованість даних в кожному окремому контексті, а також спростити визначення координат розпізнаних дефектів. При цьому низька якість даних інспекції обумовлює актуальність дослідження ефективності використання та поєднання ідей і методів композиції моделей, сіамських мереж та інформаційно-екстремальної інтелектуальної технології для задач розпізнавання функціонального стану труб водовідведення.

Основні наукові результати, наведені у першому розділі, опубліковано у працях автора [19], [22], [23], [24], [70].

РОЗДІЛ 2

РОЗРОБЛЕННЯ МОДЕЛІ ТА МЕТОДУ НАВЧАННЯ ДЛЯ ОЦІНЮВАННЯ ФУНКЦІОНАЛЬНОГО СТАНУ ТРУБ ВОДОВІДВЕДЕННЯ

2.1 Аналіз слабо формалізованого процесу оцінювання функціонального стану труб водовідведення та обґрунтування вибору методу досліджень

Мережі трубопроводів водовідведення є важливим елементом комунальної інфраструктури, який дозволяє безпечно транспортувати стічні води від домогосподарств, підприємств та дощові води з вулиць до спеціалізованих споруд. Найбільш поширеним методом оцінки стану труби водовідведення є використання дистанційно керованих, зазвичай через кабель, мобільних роботів (кравлерів). Вони, як правило, оснащені відеокамерою (ССТV) та джерелом світла. Такий робот опускається через колодязь в трубу водовідведення, оператор керує його пересуванням і направляє та фокусує камеру на підозрілих ділянках. Далі дані завантажують в хмарний сервіс, де їх аналізує оператор. За результатами аналізу здійснюється оцінювання функціонального стану труб водовідведення, що описується кодами дефектів та їх координати відповідно до прийнятих стандартів і оформлюється у вигляді звіту про інспекцію. Повздовжні координати зчитуються одометром колес, а радіальні координати визначають за годинниковою стрілкою і часто подаються в огрубленому вигляді. В багатьох Європейських стандартах використовуються слова : зліва, справа, вгорі, внизу, або вгорі зліва чи справа, або внизу зліва чи справа. Для такої процедури характерні повільність перевірки відео та наявність помилок, що зумовлене впливом суб'єктивних факторів людини-інспектора (рис. 2.1). Вказане вище обумовлює значний науковий та комерційний інтерес до автоматизації процесу перевірки даних відео інспекції через використання інтелектуальних технологій – машинного зору та машинного навчання [9, 48].

У відомих розробках щодо інтелектуального аналізу даних відео інспекції труб водовідведення не враховують просторово-часовий контекст та відсутні

достатні заходи щодо врахування значної кількості артефактів та шумів. На якість зображень відеоінспекції впливають засвічування поверхні, бризки води, тимчасові неполадки відеопристрою, дефекти зображень, схожість бруду та піни на реальні дефекти і тому подібне (рис. 2.2). При цьому неоднозначність кодування дефектів складної конфігурації, висока варіативність візуальних спостережень, рутинність і дороговартісність розмітки даних, обумовлюють обмеженість обсягу і високу незбалансованість навчальних даних. Це призводить до збільшення впливу вибору архітектури, гіперпараметрів та початкової ініціалізації вагових коефіцієнтів на ефективність навчання.

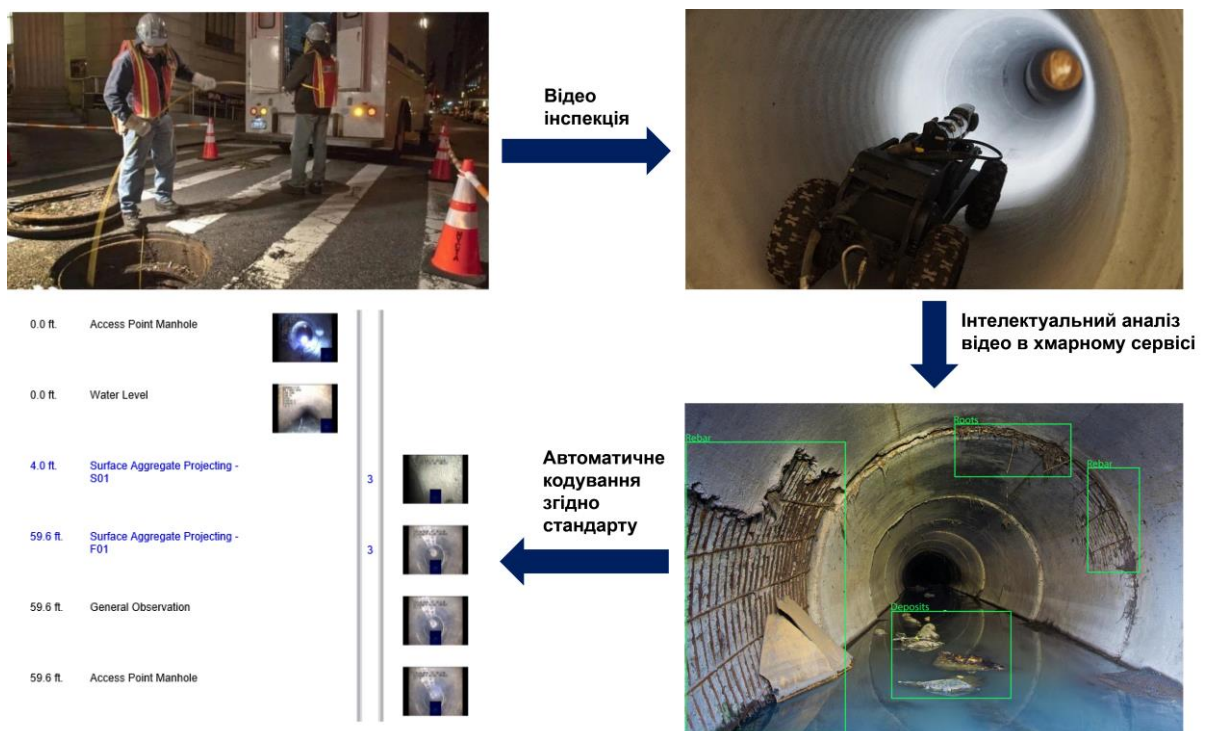


Рисунок 2.1 – Схема процесу формування звіту про інспекцію труб водовідведення

Одним з перспективних шляхів підвищення функціональної ефективності системи інтелектуального аналізу даних відеоінспекції труб водовідведення за умов ресурсних та інформаційних обмежень є використання принципів переносу знань (transfer learning) та алгоритмів настройки гіперпараметрів. Однак візуальні спостереження в інспекції труб водовідведення досить сильно відрізняються від

зображень відкритих наборів даних, на яких навчаються state-of-the-art моделі, що досить сильно знижує ефективність техніки переносу знань [71]. У таких випадках найбільш виправданим підходом є використання великого обсягу нерозмічених даних для ініціалізації моделі аналізу відеокадрів. Проте ефективність методу самонавчання для ініціалізації нейромережі не завжди передбачувана, вона залежати від особливостей задач на кінці мережі (downstream task) і потребує дослідження в кожному конкретному випадку.

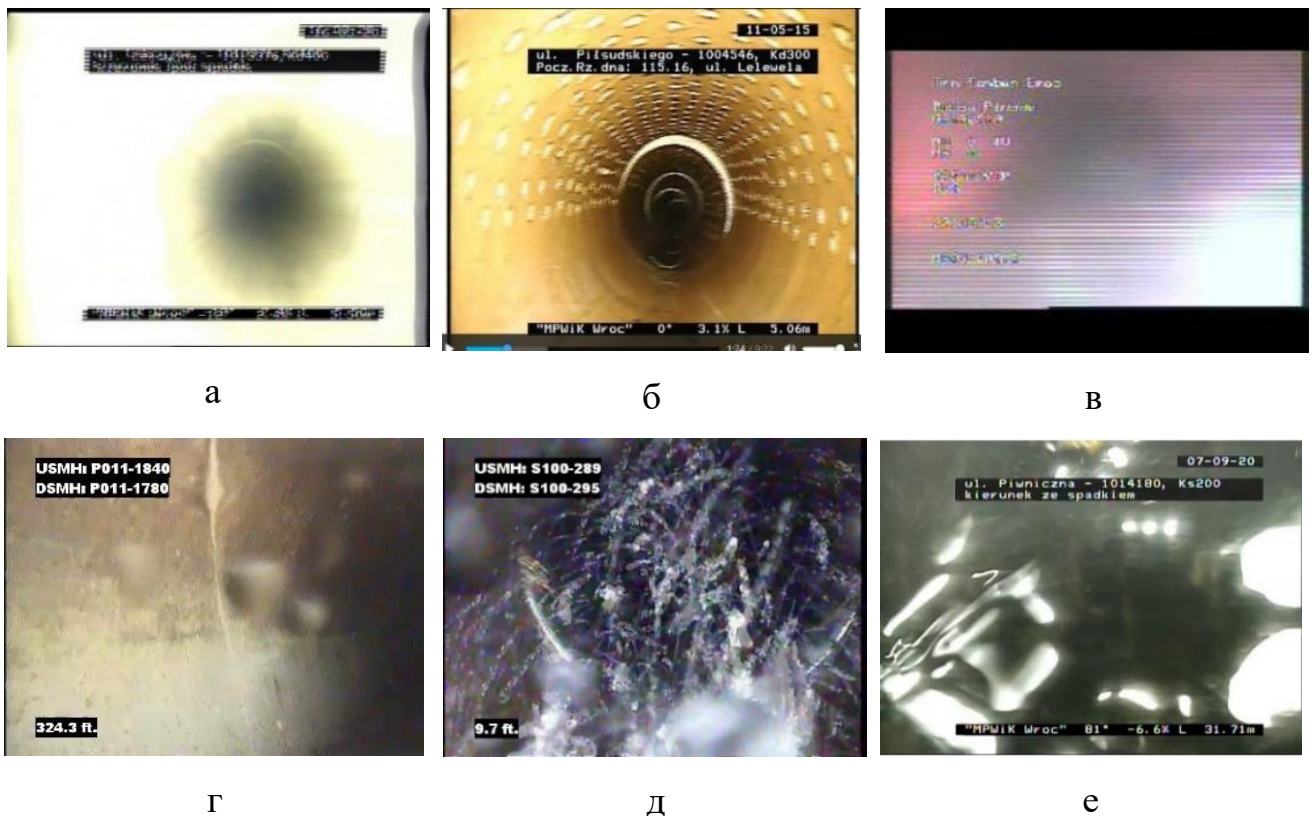


Рисунок 2.2 – Відеокадри з новизною, шумами та артефактами :
 а – засвічування поверхні; б – новизна у вигляді множини отворів;
 в – неполадки відеопристрою; г – краплини на склі;
 д – бризки води; е – бліки на воді

На підставі аналітичного огляду праць попередників, проведеного в першому розділі, було сформовано шляхи удосконалення окремих функціональних компонентів інформаційної технології аналізу даних відеоінспекції труб водовідведення, які наведені у табл. 2.1.

Таблиця 2.1 – Аналіз існуючих недоліків інформаційної технології аналізу даних відеоінспекції труб водовідведення

Складові інформаційної технології	Невирішені задачі та виявлені проблеми
<p>Екстракція інформативного ознакового опису спостережень для інтелектуального аналізу даних відеоінспекції труб водовідведення [48, 71].</p>	<p>Відсутнє використання у моделях екстракції ознакового опису контекстної просторово-часової інформації, що знижує достовірність рішень в умовах неповної інформації.</p>
<p>Ефективне машинне навчання за умов обмеженого і незбалансованого обсягу розмічених навчальних даних [31, 72].</p>	<p>Ефективність існуючих методів глибокого машинного навчання за умов обмеженого і незбалансованого обсягу розмічених даних залежить від початкової ініціалізації вагових коефіцієнтів. Прийнятні для практичного використання результати вдається отримати лише при великих і збалансованих обсягах розмічених даних.</p>
<p>Побудова оптимальних вирішувальних правил для детектування дефектів труб водовідведення [62, 73].</p>	<p>Розмітка даних для традиційних алгоритмів детектування об'єктів є рутинною і дорогавартісною. Зашумленість і незбалансованість вибіркового даних та висока варіативність візуальних спостережень в задачі аналізу даних інспекції труб водовідведення обумовлює необхідність калібрування порогів спрацювання. Однак подібне калібрування ефективна лише для класів розпізнавання, що мають великий обсяг розмічених зразків.</p>

Продовження таблиці 2.1

Складові інформаційної технології	Невирішені задачі та виявлені проблеми
Оцінювання ефективності екстрактора ознакового опису та вирішувальних правил в режимі навчання та екзамену [74, 75].	Традиційні критерії оптимізації параметрів і гіперпараметрів не забезпечують реалізацію інформаційного пляшкового горла, тому не забезпечують оптимальної робастності до шуму і незабалансованості навчальних даних.
Реалізація інформаційної системи аналізу даних відеоінспекції труб водовідведення [76, 77].	Забезпечення прийнятної достовірності рішень і рентабельність впровадження системи інтелектуального аналізу знаходяться у протиріччі, внаслідок високих витрат на створення і експлуатацію.

З табл. 2.1 видно, що існуючі моделі аналізу даних відеоінспекції потребують удосконалення шляхом врахування як орієнтації камери відеоспостереження, так і міжкадрових залежностей для моделювання просторово-часового контексту. При цьому методи навчання повинні бути модифіковані для ефективного використання як розмічених так і нерозмічених даних.

Детальне вивчення задачі кодування функціонального стану труб системи водовідведення в рамках існуючих стандартів показує, що вона може бути вирішена на основі моделей класифікаційного аналізу зображень. Це спрощує процес розмітки даних порівняно з задачею детектування зон чи об'єктів інтересу.

Для оцінювання орієнтації камери достатньо використати класифікатор зображень. Назвемо його класифікатором контекстів, оскільки в алфавіт класів можуть бути включені різноманітні нестандартні ситуації, які необхідно правильно обробляти. На рис. 2.3 показано зразки класів, що можуть бути включені до алфавіту класів класифікатора контекстів. У випадку врахування сусідніх кадрів

алфавіт може трансформуватися шляхом заміни малоінформативних контекстів на більш інформативні.



Рисунок 2.3 – Приклади зразків кожного класу контекстів спостереження:
 а – клас “орієнтація вперед”; б – клас “напівповорот вліво”; в – клас “напівповорот вправо”; г – клас “напівповорот вгору”; г – клас “напівповорот вниз”; д – клас “впритул до стінки”; е – клас “колодязь”; є – клас “ігнорувати”; ж – клас “руйнування труби”; з – клас “під’єднання до труби водовідведення”

Наприклад, класи “напівповороту ” та “ігнорувати” можна прибрати, а клас “під’єднання до труби водовідведення” розширити класами “під’єднання до труби водовідведення зліва”, “під’єднання до труби водовідведення справа”,

“під’єднання до труби водовідведення вгори” та “під’єднання до труби водовідведення вниз”. Класи “напівповорот вліво”, “напівповорот вправо”, “напівповорот вгору” та “напівповорот вниз” можуть бути уточнені до класів “зліва”, “справа”, “вгори” та “внизу” відповідно.

Переважна більшість протоколів інспекції передбачає, що оператор кравлера повертає камеру для детальнішого огляду потенційного дефекту. Для зменшення варіативності даних внаслідок поворотів камери, можна окремо розглядати задачі розпізнавання в контексті, де камера орієнтована до стінки труби, і окремо в контексті, де камера орієнтована вперед вздовж труби [78, 79, 80].

У випадку камери орієнтованої вперед вздовж труби важливо не розпізнавати віддалені дефекти, для яких поточне значення одометра не є актуальним [78]. Тому аналіз зображення варто здійснювати тільки в областях, прилеглих до його меж. На рис. 2.4 показано зразки класів дефектів і виділено червоним бокові патчі з актуальним значення одометра і наявністю дефекту.

У випадку контекстів, що відповідають орієнтації камери вбік до стінки труби, алфавіт класів дефектів може бути розширений класами “відкритий стик”, “зміщений стик”, “пошкодження поверхні”, “осад” та “порушення цілісності прокладки на стику”. А клас “корені” розщеплений на класи “тоненькі корені” та “значне розростання коренів [79]. При цьому аналізу можна піддавати все зображення, оскільки всі його ділянки відповідатимуть актуальному значенню одометра. Для розпізнавання просідання труби внаслідок підмивання чи значного навантаження потрібно відслідковувати зміну рівня води, яку так само можна розпізнавати класифікатором.

Як було показано в першому розділі розроблення алгоритмів синтезу моделей класифікаційного аналізу дефектів труб водовідведення перспективно здійснювати з використанням ідей і методів сіамських мереж, контрастного самонавчання, завадозахищеного кодування та інформаційно-екстремальної інтелектуальної технології. В рамках даних підходів існує принципова можливість побудови високодостовірної моделі аналізу даних за умов високої варіативності, зашумленості та обмеженого обсягу розмічених навчальних даних. Проте питання

ефективної інтеграції відповідних методів і вибору архітектури моделей для класифікаційного аналізу даних відеоінспекції труб водовідведення є недостатньо дослідженим.

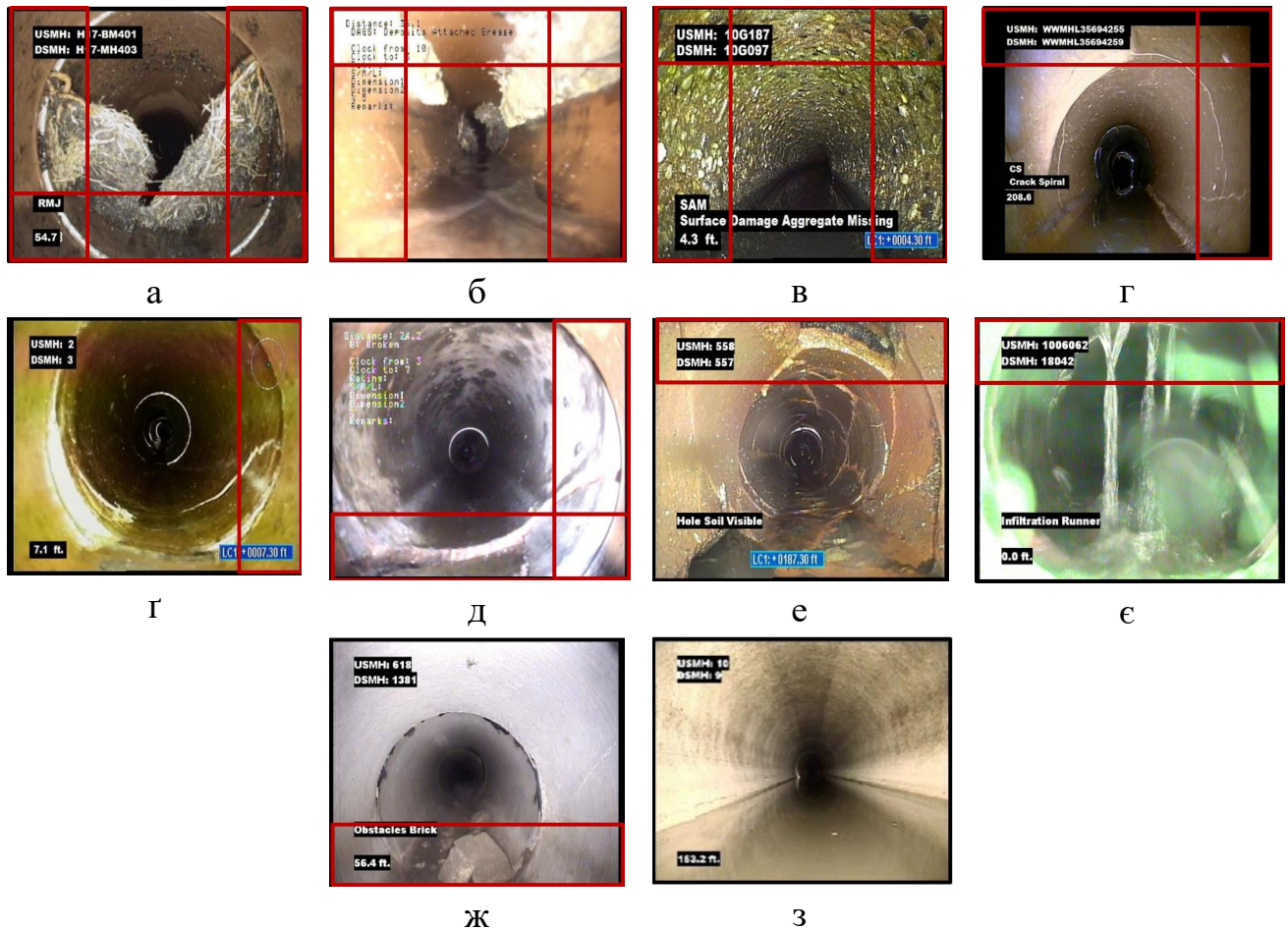


Рисунок 2.4 – Приклади зразків кожного класу дефектів труб :

- а – клас “корені”;
- б – клас “відкладення жиру чи солі”;
- в – клас “пошкодження поверхні”;
- г – клас “тріщина”;
- д – клас “розлом”;
- е – клас “дірка в трубі”;
- е – клас “інфільтрація”;
- ж – клас “перешкода в трубі”;
- з – нормальний стан

Вибір напрямку дисертаційних досліджень потребує розв’язання низки науково-методологічних задач, пов’язаних з аналізом і синтезом моделі оцінювання функціонального стану труб водовідведення за даними відеоінспекції, таких як:

1) розроблення моделі екстракції ознакового опису для врахування просторово-часового контексту відеоспостережень, що функціонує за умов неповної визначеності;

2) розроблення моделі розпізнавання дефектів труб водовідведення;

3) розроблення моделі розпізнавання рівня води в трубах водовідведення;

4) розроблення методу машинного навчання для формування інформативного ознакового опису спостережень за умов обмеженого і незбалансованого обсягу розмічених зразків;

5) розроблення методу побудови вирішувальних правил для моделі оцінювання функціонального стану труб водовідведення;

6) розроблення інформаційної технології для синтезу інтелектуальної системи оцінювання функціонального стану труб водовідведення за даними відеоінспекції.

Таким чином, аналіз існуючих рішень для реалізації інтелектуальної інформаційної технології оцінювання функціонального стану труб водовідведення показав необхідність підвищення їх ефективності для використання в умовах неповної визначеності. При цьому обґрунтовано доцільність використання як напрямку дисертаційних досліджень ідей і методів нейромережевого контрастного навчання та інформаційно-екстремальної інтелектуальної технології аналізу даних.

2.2 Формалізована постановка задачі

Розглянемо формалізовану постановку задачі створення інтелектуальної інформаційної технології аналізу даних для оцінювання функціонального стану трубопроводів водовідведення за даними відеоінспекції. При цьому функціональний стан труб водовідведення визначається інформацією про типи дефектів і їх розміщення. Нехай дано : $\{X_s^0 | s = \overline{1, S}\}$ – алфавіт класів розпізнавання, що характеризує S контекстів спостереження в трубі водовідведення, частина з яких стосується орієнтації камери; $\{X_u^0 | u = \overline{1, U}\}$ –

алфавіт класів розпізнавання, що характеризує дефекти труби водовідведення;

$D_{\text{unlabeled}} = \{ x_j \mid j = \overline{1, n} \}$, $D_{\text{context}}^{\text{train}} = \{ x_{c,j} \mid c = \overline{1, C}; j = \overline{1, n_c^{\text{train}}} \}$, $D_{\text{context}}^{\text{test}} = \{ x_{c,j} \mid c = \overline{1, C}; j = \overline{1, n_c^{\text{test}}} \}$ та $D_{\text{defect}}^{\text{train}} = \{ x_{m,j} \mid m = \overline{1, M}; j = \overline{1, n_m^{\text{train}}} \}$, $D_{\text{defect}}^{\text{test}} = \{ x_{m,j} \mid m = \overline{1, M}; j = \overline{1, n_m^{\text{test}}} \}$, а також $D_{\text{water}}^{\text{train}} = \{ x_{w,j} \mid w = \overline{1, W}; j = \overline{1, n_w^{\text{train}}} \}$ і $D_{\text{water}}^{\text{test}} = \{ x_{w,j} \mid w = \overline{1, W}; j = \overline{1, n_w^{\text{test}}} \}$ – набори нерозмічених та розмічених навчальних та тестових даних, C – потужність алфавіту класів контексту спостереження, M – потужність алфавіту класів, що характеризують дефекти труби водовідведення, W – потужність алфавіту класів, що характеризують рівень води в трубі. При цьому n , n_c^{train} , n_c^{test} , n_m^{train} , n_m^{test} , n_w^{train} , n_w^{test} – обсяги нерозмічених та розмічених навчальних та тестових вибірок відповідних класів.

Дано структурований вектор просторово-часових параметрів комплексної моделі аналізу даних відеоінспекції, який у загальному випадку має структуру

$$g = \langle e_1, \dots, e_{\xi_1}, \dots, e_{\Xi_1}, f_1, \dots, f_{\xi_2}, \dots, f_{\Xi_2} \rangle, \quad \Xi_1 + \Xi_2 = \Xi, \quad (2.1.1)$$

де $e_1, \dots, e_{\xi_1}, \dots, e_{\Xi_1}$ – параметри, які прямо впливають на екстракцію ознакового опису спостережень;

$f_1, \dots, f_{\xi_2}, \dots, f_{\Xi_2}$ – параметри, які прямо впливають на ефективність класифікатора спостережень.

При цьому відомі обмеження на параметри функціонування :

$$\begin{aligned} R_{\xi_1}(e_1, \dots, e_{\xi_1}, \dots, e_{\Xi_1}) &\leq 0, \\ R_{\xi_2}(f_1, \dots, f_{\xi_2}, \dots, f_{\Xi_2}) &\leq 0. \end{aligned} \quad (2.1.2)$$

Необхідно

1. На етапі машинного навчання інтелектуальної моделі аналізу даних визначити оптимальний вектор параметрів, що забезпечує максимум критерію ефективності

$$g^* = \arg \max_G \{ J(g) \} \quad (2.1.3)$$

де J – критерій ефективності для всієї моделі аналізу даних, що оцінюється на тестовій вибірці; G – допустима область значень параметрів, які впливають на екстракцію ознак і прийняття рішень.

2. Розробити модель і метод інформаційного синтезу екстрактора ознакового опису спостережень для інтелектуального аналізу даних відеоінспекції труб водовідведення, який повинен відповідати таким основним вимогам :

- враховувати просторово-часовий контекст спостережень для підвищення інформативності ознакового опису і зменшення невизначеності внаслідок шуму, артефактів та збоїв відеообладнання;

- забезпечувати високу функціональну ефективність навчання екстрактора ознакового опису для задач класифікаційного аналізу кадрів відеоспостереження за умов високої варіативності та зашумленості спостережень, а також обмеженого обсягу розмічених навчальних даних, що має місце на практиці;

- реалізувати ініціалізацію вагових коефіцієнтів моделі екстрактора ознак на основі самонавчання для ефективного використання нерозмічених навчальних зразків, яких на багато більше ніж розмічених;

3. На етапі екзамену, тобто безпосередньо у робочому режимі, необхідно за максимумом функції належності спостереження до одного з класів із заданого алфавіту розпізнати поточний функціональний стан на t -му кадрі відеоінспекції

$$\langle k^* \rangle = \underset{G_\mu}{\operatorname{arg\,max}} \{ \mu_k(f(x_t)) \}, \quad k = \overline{1, K}, \quad (2.1.4)$$

де $\mu_m(\bullet)$ – значення функції належності спостереження до k -го класу із алфавіту.

Таким чином, задача створення інтелектуальної інформаційної технології аналізу даних для оцінювання функціонального стану трубопроводів водовідведення за даними відеоінспекції полягає в оптимізації параметрів моделі з метою забезпечення максимуму критерію ефективності моделі, що оцінюється на вибірковій множині даних.

2.3 Моделі відображення множин для опису процесу навчання

У загальному випадку процес машинного навчання полягає у оптимізації параметрів моделі для трансформації простору ознак і побудови моделі для прийняття рішень на основі високорівневого ознакового опису спостережень. Даний процес можна описати у вигляді діаграми відображення множин, задіяних у процесі аналізу даних (рис. 2.5) [66].

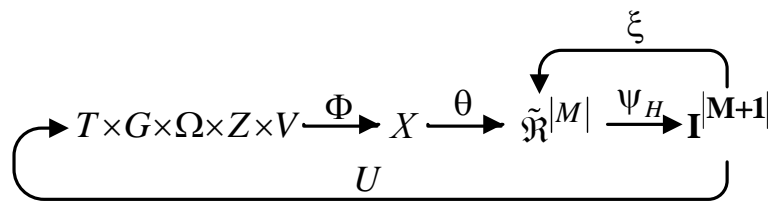


Рисунок 2.5 – Узагальнена модель процесу навчання у вигляді діаграми відображення множин

На рис. 2.5 прийнято такі позначення :

T – множина моментів часу зчитування інформації;

G – факторів, які впливають на об’єкт спостереження;

Ω – простір первинних ознак розпізнавання, наприклад, яскравість пікселів кожного каналу зображення заданої роздільної здатності;

Z – простір можливих функціональних станів (класів), у яких перебуває спостережуваний об’єкт чи процес;

V – множина типів вирішувальних правил;

X – вибіркова множина (вхідна навчальна вибірка);

$\Phi: G \times T \times \Omega \times Z \times V \rightarrow X$ – оператор формування вибіркової навчальної множини X ;

θ – оператор відображення вибіркової множини X в простір вторинних ознак, де сформовано M вирішувальних правил $\tilde{\mathfrak{R}}^{|M|}$ ($\tilde{\mathfrak{R}}^{|M|} \subset \Omega$);

ψ – оператор прийняття рішень, наприклад, рішень щодо належності поточного функціонального стану до одного з класів розпізнавання;

$I^{|M+1|}$ – множина допустимих гіпотез (рішень), де γ_{M+1} – гіпотеза про відмову прийняття рішення;

ξ – оператор зворотного зв'язку;

$U : I^{|M+1|} \rightarrow T \times G \times \Omega \times Z \times V$ – оператор регламентації процесів навчання та екзамену.

На рис. 2.5 оператор зворотного зв'язку ξ корегує геометричні параметри вирішувальних правил $\tilde{\mathfrak{R}}^{|M|}$ з метою максимізації точнісних характеристик моделі аналізу даних. Спосіб реалізації даного оператора визначається підходом і методологією синтезу інтелектуального алгоритму аналізу даних.

В задачах відеоінспекції труб водовідведення існує потреба у розпізнаванні контексту, рівня води, наявності дефектів, типу і розміщення дефектів. Задача визначення розміщення дефектів традиційно є задачею регресійного аналізу, однак її можна звузити до задачі класифікаційного аналізу за умов доступності інформації про орієнтацію камери відеоспостереження (рис. 2.4). Тобто задача аналізу даних відеоінспекції труб водовідведення може бути декомпонована до трьох окремих задач класифікаційного аналізу, результатів вирішення яких достатньо для кодування функціонального стану труби згідно існуючих стандартів [79, 80, 81]:

- класифікаційний аналіз контекстів спостереження, до алфавіту класів яких належать різні орієнтації камери та стани початку і закінчення інспекції;
- класифікаційний аналіз фіксованих ділянок зображення на класи дефектів або нормальний функціональний стан;
- класи рівнів води в трубопроводі для виявлення ситуацій просідання труби під зовнішнім тиском чи внаслідок підмивання ґрунту.

Оскільки точна розмітка даних вимагає узгодження від декількох ліцензованих операторів для зменшення впливу суб'єктивних факторів, то існує проблема обмеженого і незбалансованого обсягу розмічених навчальних даних, що підвищує чутливість глибоких нейромережових алгоритмів до початкової ініціалізації вагових коефіцієнтів. При цьому великі обсяги нерозмічених даних, або даних з неузгодженою розміткою є доступними, проте, як правило, не

використовуються. Крім того, для підвищення узагальнюючої здатності та завадозахищеності алгоритмів класифікаційного аналізу даних найбільш простим і ефективним методом вважається кодування міток класів завадозахищеними двійковими кодами, наприклад з таблиці Адамара [68]. Однак для підвищення гладкості моделі аналізу даних і врахування внутрішньої структури даних доцільно використовувати методи вбудовування міток (Label Embedding), що може бути реалізовано за рахунок поєднання різних технік глибокого хешування (Deep Hashing) та адаптивного кодування цільових міток (Target Coding) [82]. Найпростішим методом узгодження цих методів і підходів є реалізація навчання моделі аналізу даних в декілька послідовних етапів, де кожен попередній етап здійснює ініціалізацію вагових коефіцієнтів екстрактора ознак для наступного етапу. При цьому окремі етапи можуть повторюватися для реалізації принципу неостаточних рішень Д. Габора [83]. На рис. 2.6 показано діаграму відображення множин, задіяних під час навчання моделі для класифікаційного аналізу даних.

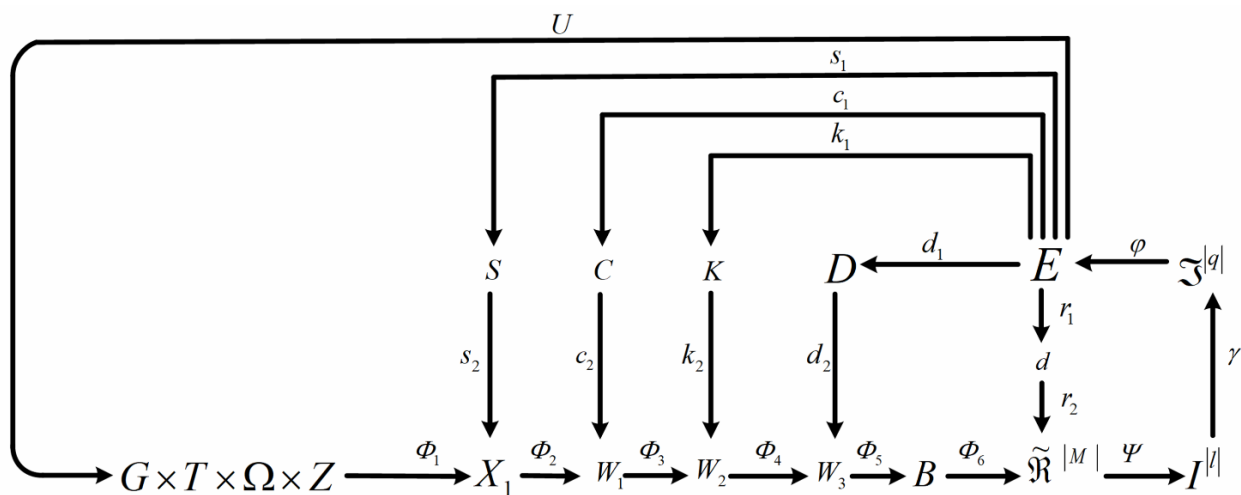


Рисунок 2.6 – Діаграма відображення множин, задіяних під час класифікаційного аналізу даних відеоінспекції труб водовідведення

На рис. 2.6 прийнято такі нові позначення :

B – спостереження, закодовані двійковим кодом в результаті навчання екстрактора високорівневих ознак;

E – множина значень критерію ефективності вирішувальних правил при

різних значеннях параметрів функціонування;

$\tilde{\mathfrak{R}}^{|M|}$ – розбиття на M класів у двійковому просторі високорівневих вторинних ознак (наприклад, простір Хеммінга);

$I^{|l|}$ – множина гіпотез при l -альтернативному рішенні;

$\mathfrak{S}^{|q|}$ – множина точнісних характеристик рішень;

S, C, K, D – множини гіпер-параметрів для алгоритмів навчання з самоучителем для утилізації нерозмічених даних, навчання з учителем для кодування класів, навчання з учителем з завадозахищеними двійковими цільовими мітками, значень порогів бінаризації ознак відповідно;

$s_1 \circ s_2, c_1 \circ c_2, d_1 \circ d_2$ – контури навчання екстрактора ознак з самоучителем, навчання з учителем для кодування класів, навчання з учителем для точного настроювання екстрактора ознак з фіксованими кодами класів;

$r_1 \circ r_2$ – контур ітеративного відновлення оптимального розбиття простору високорівневих ознак на M класів розпізнавання;

$\Phi_1 : G \times T \times \Omega \times Z \rightarrow X_1$ – оператор формування вибіркової множини даних;

$\Phi_2 : X_1 \rightarrow W_1$ – оператор навчання з самоучителем, що оновлює параметри екстрактора ознак;

$\Phi_3 : W_1 \rightarrow W_2$ – оператор навчання з учителем для визначення оптимального кодування класів;

$\Phi_4 : W_2 \rightarrow W_3$ – навчання з учителем з цільовими мітками класів, закодованих двійковими кодами, для уточнення параметрів екстрактора ознакового опису;

$\Phi_5 : W_3 \rightarrow B$ – оператор обчислення двійкової навчальної матриці;

$\Phi_6 : B \rightarrow \tilde{\mathfrak{R}}^{|M|}$ – оператор розбиття простору ознак на класи еквівалентності.

На рис. 2.7 наведено відповідну функціональну схему моделі навчання.

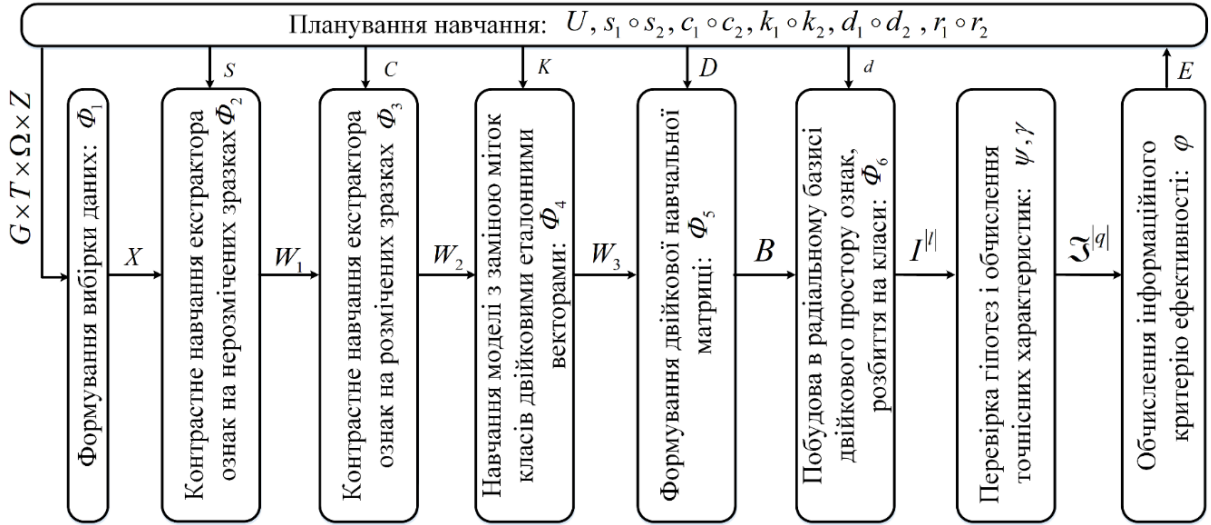


Рисунок 2.7 – Функціональна схема моделі інтелектуального аналізу даних відеоінспекції труб водовідведення

Процес розпізнавання функціонального стану з використанням навченої моделі аналізу даних можна описати діаграмою відображення множин, що показана на рис. 2.8 [66].

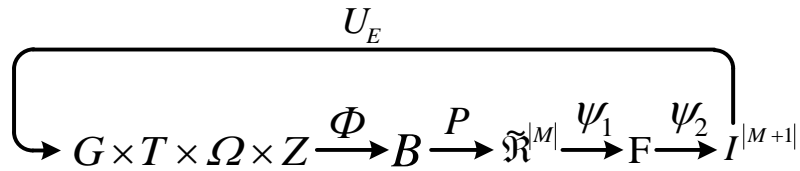


Рисунок 2.8 – Діаграма відображення множин для опису процесу розпізнавання функціонального стану

У діаграмі на рис. 2.8 оператор $\Phi = \Phi_1 \circ \Phi_2 \circ \Phi_3 \circ \Phi_4 \circ \Phi_5$ відображає універсум випробувань на вибіркочну множину B . Вибіркочну множину B можна представити у вигляді двійкової матриці $\|b_i^{(j)}\|_{i=\overline{1, N}, j=\overline{1, n}}$, що формується тими ж операторами трансформації простору ознак, що використовувалися під час машинного навчання, за винятком того, що параметри екстракції ознак фіксовані. Оператор P відображає вибіркочну двійкову множину $B \subset \Omega_B$ на розбиття $\mathfrak{R}^{|M|}$, що

оптимізоване в режимі навчання [85]. Оператор класифікації ψ утворює композицію двох операторів $\psi = \psi_1 \circ \psi_2$. Оператор ψ_1 обчислює значення функції належності до кожного з класів розпізнавання: $\psi_1: \mathfrak{R}^{|M|} \rightarrow F$, де F – множина функцій належності. Оператор ψ_2 вибирає гіпотезу за максимальним значенням функції належності $\psi_2: F \rightarrow I^{|M+1|}$ [84].

Моделі інтелектуального аналізу даних у вигляді діаграм відображення множин набули найбільшого поширення в рамках інформаційно-екстремальної інтелектуальної технології (ІЕІ-технології) [66]. Такий підхід до моделювання забезпечує високорівневе подання інтелектуального алгоритму і дозволяє встановлювати відношення між елементами інформаційного забезпечення та інформаційними потоками оброблення даних. Моделі відображення множин суттєво полегшують реінжиніринг та рефакторинг інформаційного забезпечення створеної інтелектуальної системи.

Таким чином, запропоновані моделі відображення множин та функціональні схеми для опису алгоритмів навчання та екзамену моделі класифікаційного аналізу даних відеоінспекції дозволяють виконати кодування вхідних спостережень у вигляді інформативного ознакового подання, що зручне для побудови простих вирішувальних правил. При цьому реалізація багатоетапного методу машинного навчання і впровадження відповідних контурів оптимізації гіперпараметрів алгоритмів навчання з самоучителем та учителем забезпечує ефективне використання як розмічених так і нерозмічених навчальних даних.

2.4 Критерії оцінювання ефективності моделі класифікаційного аналізу даних

Оцінку ефективності моделі класифікаційного аналізу на кожному етапі її синтезу здійснюють згідно принципу зовнішнього доповнення як на навчальній розміченій вибірці, так і на зовнішній незалежній розміченій вибірці, яку називають тестовою або перевірковою [83]. На практиці для оцінювання функціональної

ефективності класифікатора можуть мати місце такі допущення :

- рішення є двохальтернативним, оскільки задача багатокласової класифікації може бути зведена до серії двохкласових;
- рівноймовірність функціональних станів в умовах невизначеності згідно принципу Бернуллі-Лапласа.

У випадку кодування класів двійковими кодами для кожного класу в просторі високорівневих ознак будується замкнута роздільна гіперповерхня – гіперсферичний (або псевдогіперсферичний) контейнер класу розпізнавання [24]. Критерії ефективності класифікатора обчислюються на основі підрахунку результатів статистичних тестів, тобто перевірки гіпотез. Якщо появу спостереження, що належить першому класу позначити як подію μ_1 , а появу спостереження альтернативного класу як μ_2 , то оцінювання побудованого контейнера класу зводиться до перевірки основної гіпотези γ_1 про потрапляння спостереження у контейнер першого класу та альтернативної гіпотези про непотрапляння вхідного зразка до даного контейнера при появі μ_1 та μ_2 . Для кожного вхідного спостереження можливі чотири результати статистичного тесту (рис. 2.9) :

1) істинно позитивний (True Positive) – якщо спостереження 1-го класу знаходиться всередині контейнера даного класу, тобто якщо зразок 1-го класу розпізнається як спостереження 1-го класу;

2) хибно негативний (False Negative) – якщо спостереження 1-го класу знаходиться поза межами контейнера даного класу, тобто якщо зразок 1-го класу не розпізнається як спостереження 1-го класу;

3) істинно негативний (True Negative) – якщо спостереження альтернативного класу (2-го класу) знаходиться поза межами контейнера 1-го класу, тобто якщо зразок альтернативного класу не розпізнається як спостереження 1-го класу;

4) хибно позитивні (False Positive) – якщо спостереження альтернативного класу (2-го класу) знаходиться всередині контейнера 1-го класу, тобто якщо зразок

альтернативного класу розпізнається як спостереження 1-го класу.

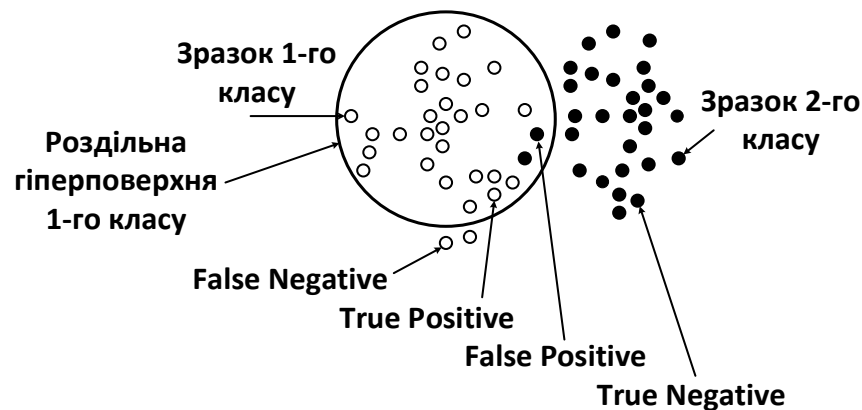


Рисунок 2.9 – Ілюстрація результатів статистичного тесту для окремих зразків при заданій роздільній гіперповерхні

Для підрахунку істинно позитивних, хибно негативних, істинно негативних та хибно позитивних результатів вводять відповідні лічильники : TP, FN, TN, FP.

Популярними характеристиками класифікаційної моделі аналізу даних вважаються чутливість, специфічність, точність, частоту помилок першого та другого роду, прецизійність та F-міру. На основі цих характеристик зручно порівнювати результати різних авторів та результати крос-валідації під час оптимізації гіперпараметрів [13].

Чутливість (Sensitivity, або True Positive Rate або Recall) – кількість істинно позитивних результатів, поділена на реальну кількість позитивних зразків. Чутливість для контейнера 1-го класу, D_1 , обчислюється за формулою

$$D_1 = p(\gamma_1 / \mu_1) = \text{Чутливість} = \frac{TP}{TP + FN}. \quad (2.4.1)$$

Специфічність (або True Negative Rate) – кількість істинно негативних результатів, поділена на реальну кількість негативних зразків. Специфічність для контейнера m -го класу, D_2 , обчислюється за формулою

$$D_2 = p(\gamma_2 / \mu_2) = \text{Специфічність} = \frac{TN}{TN + FP}. \quad (2.4.2)$$

Оцінка ймовірності помилок першого роду (або False Negative Rate) –

кількість хибно негативних результатів, поділена на загальну кількість істинно позитивних зразків. Оцінка ймовірності помилок першого роду для контейнера 1-го класу, α , обчислюється за формулою

$$\alpha = p(\gamma_2 / \mu_1) = \text{Оцінка ймовірності пропусків} = \frac{FN}{FN + TP}. \quad (2.4.3)$$

Оцінка ймовірності помилок другого роду (або False Positive Rate) – кількість хибно позитивних результатів поділена на загальну кількість істинно негативних зразків. Оцінка ймовірності помилок другого роду для контейнера 1-го класу, β , обчислюється за формулою

$$\beta = p(\gamma_1 / \mu_2) = \text{Оцінка ймовірності хибних спрацювань} = \frac{FP}{FP + TN}. \quad (2.4.4)$$

Прецизійність (або Precision) – кількість істинно позитивних результатів поділена на загальну кількість позитивних прогнозів. Прецизійність для контейнера m -го класу обчислюється за формулою

$$\text{Прецизійність} = \frac{TP}{TP + FP}. \quad (2.4.5)$$

Точність (або Accuracy) – частка правильних прогнозів від загальної кількості прогнозів. Точність для контейнера m -го класу обчислюється за формулою

$$\text{Точність} = \frac{TP + TN}{TP + TN + FP + FN}. \quad (2.4.6)$$

На практиці найбільшій увазі приділяють чутливості та прецизійності. Чим вони вищі тим ефективніша класифікаційна система, однак ці характеристики не можуть бути максимізовані одночасно, тому доводиться знаходити між ними деякий баланс. Для поєднання чутливості та прецизійності у збалансованій формі в єдиній метриці використовують F-міру (F1), що є гармонічним середнім між чутливістю і прецизійністю [79]

$$F1 = \frac{2}{\frac{1}{\text{Прецизійність}} + \frac{1}{\text{Чутливість}}} = \frac{2TP}{2TP + FP + FN}. \quad (2.4.7)$$

Для оптимізації геометричних параметрів розбиття простору ознак на еквівалентність класів в рамках ІЕІ-технології прийнято використовувати

логарифмічні інформаційні критерії ефективності. Дані критерії забезпечують побудову робастних вирішувальних правил за найскладніших у статистичному сенсі умов і прямо вимірюють рівень зняття невизначеності. Найбільш теоретично обгрунтованою в рамках ІЕІ-технології є нормалізована модифікація ентропійної міри Шенона, що обчислюється за формулою [19]

$$E = \frac{H_0 - H(\gamma)}{H_0}, \quad (2.4.8)$$

де H_0 – апіорна (безумовна) ентропія для M функціональних станів (класів) і відповідної їм множини гіпотез $\{\gamma_l | l = \overline{1, M}\}$,

$$H_0 = -\sum_{l=1}^M p(\gamma_l) \log_2 p(\gamma_l); \quad (2.4.9)$$

$H(\gamma)$ – апостеріорна умовна ентропія, яка характеризує залишкову невизначеність після прийняття рішень

$$H(\gamma) = -\sum_{l=1}^M p(\gamma_l) \sum_{m=1}^M p(\mu_m / \gamma_l) \log_2 p(\mu_m / \gamma_l) \quad (2.4.10)$$

де $p(\gamma_l)$ – апіорна ймовірність прийняття гіпотези $p(\mu_m / \gamma_l)_l$;

$p(\mu_m / \gamma_l)$ – апостеріорна ймовірність появи події μ_m за умови прийняття гіпотези γ_l ;

M – число альтернативних гіпотез.

З урахуванням двохальтернативності і рівноймовірності рішень ($p(\gamma_1) = p(\gamma_2) = 0,5$) та формул (2.4.9) і (2.4.10) критерій (2.4.8) можна записати в наступному вигляді

$$E = 1 + \frac{1}{2} \sum_{l=1}^2 \sum_{m=1}^2 p(\mu_m / \gamma_l) \log_2 p(\mu_m / \gamma_l). \quad (2.4.11)$$

Якщо використати формулу Байеса [19] та формули (2.4.1)-(2.4.4), то апостеріорні ймовірності $p(\mu_m / \gamma_l)$ можна виразити через апіорні :

$$p(\mu_1 / \gamma_1) = \frac{p(\gamma_1 / \mu_1)}{p(\gamma_1 / \mu_1) + p(\gamma_1 / \mu_2)} = \frac{D_1}{D_1 + \beta} \quad (2.4.12)$$

$$p(\mu_2 / \gamma_1) = \frac{p(\gamma_1 / \mu_2)}{p(\gamma_1 / \mu_1) + p(\gamma_1 / \mu_2)} = \frac{\beta}{D_1 + \beta} \quad (2.4.13)$$

$$p(\mu_1 / \gamma_2) = \frac{p(\gamma_2 / \mu_1)}{p(\gamma_2 / \mu_1) + p(\gamma_2 / \mu_2)} = \frac{\alpha}{D_2 + \alpha} \quad (2.4.14)$$

$$p(\mu_2 / \gamma_2) = \frac{p(\gamma_2 / \mu_2)}{p(\gamma_2 / \mu_1) + p(\gamma_2 / \mu_2)} = \frac{D_2}{D_2 + \alpha} \quad (2.4.15)$$

Після підстановки (2.4.12)-(2.4.15) в (2.4.11) отримаємо формулу для обчислення ентропійного критерію ефективності за Шенноном

$$E = 1 + \frac{1}{2} \left(\frac{\alpha}{\alpha + D_2} \log_2 \frac{\alpha}{\alpha + D_2} + \frac{D_1}{\beta + D_1} \log_2 \frac{D_1}{\beta + D_1} + \frac{\beta}{\beta + D_1} \log_2 \frac{\beta}{\beta + D_1} + \frac{D_2}{\alpha + D_2} \log_2 \frac{D_2}{\alpha + D_2} \right) \quad (2.4.16)$$

Якщо обсяг вибірок першого класу і сусіднього до нього другого класу рівні між собою, то критерій (2.4.16) можна виразити через лічильники результатів статистичних тестів [14]

$$E = 1 + \frac{1}{2} \left(\frac{FN}{FN + TN} \log_2 \frac{FN}{FN + TN} + \frac{TP}{TP + FP} \log_2 \frac{TP}{TP + FP} + \frac{FP}{FP + TP} \log_2 \frac{FP}{FP + TP} + \frac{TN}{TN + FN} \log_2 \frac{TN}{TN + FN} \right) \quad (2.4.17)$$

Під час обчислення інформаційного критерію (2.4.17) для уникнення невизначеності у першого множника кожного доданку контролюється значення чисельника. Якщо чисельник рівний нулю, то весь доданок відкидається (прирівнюється нулю) і далі даний доданок не обчислюється.

Якщо прийняти $\alpha = 1 - D_1$ та $\beta = 1 - D_2$, то можна побудувати залежність критерію (2.4.16) від чутливості і специфічності, тобто $E = f(D_1, D_2)$. Ця залежність матиме симетричний вигляд відносно площини $D_1 = D_2$, тому для уникнення невизначеності викликані неоднозначністю даної залежності вводиться робоча область інформаційного критерію у вигляді системи нерівностей

$$\begin{cases} D_1 \geq 0,5; \\ D_2 \geq 0,5. \end{cases} \quad (2.4.18)$$

У випадку багатокласової класифікації оптимізація роздільних гіперповерхонь для кожного класу відбувається окремо і у загальному випадку вибірка сусіднього класу може формуватися на основі різних стратегій :

- 1) «кожний проти всіх» – зразками сусіднього класу вважаються всі зразки решти класів;
- 2) «кожний проти найближчого сусіда» – зразками сусіднього класу вважаються зразки одного з класів, що найближчий до даного класу;
- 3) «кожний проти кожного» – побудова $M(M-1)/2$ класифікаторів, що розрізняють всі пари класів;
- 4) дихотомія – побудова двійкового дерева, в якому $M-1$ нетермінальним вузлам відповідають двохальтернативні класифікатори, а M термінальним – класи.

Перша стратегія призводить до незбалансованості класів, а друга стратегія потребує пошуку найближчого сусіднього класу, а решта стратегій ускладнюють структуру класифікатора. Тому пропонується модифікація принципу «кожний проти всіх», де для m -го класу формується вибірка з n_m найближчих зразків решти класів, де n_m – обсяг вибірки m -го класу.

Для обчислення будь-якого усередненого критерію ефективності багатокласового класифікатора можна використати мікро- або макро- усереднення. Під час обчислення макро усередненого значення критерію (метрики) значення даного критерію обчислюється незалежно для кожного класу і потім усереднюється по алфавіту класів, а під час обчислення мікро-усередненого значення критерію агрегується внесок всіх класів. У випадку незбалансованості класів мікроусереднення критеріїв є більш чутливим до зменшення функціональної ефективності класифікатора, тому слід віддавати перевагу мікроусередненню. Якщо позначити метрику продуктивності класифікатора як функцію $P()$, аргументами якої є лічильники результатів статистичного тесту класифікатора, то

значення мікроусереднення P_{micro} та макроусереднення P_{macro} для алфавіту з M класів можна розрахувати за формулами [81]:

$$P_{\text{micro}} = P \left(\sum_{m=1}^M TP_m, \sum_{m=1}^M TN_m, \sum_{m=1}^M FP_m, \sum_{m=1}^M FN_m \right) \quad (2.4.19)$$

$$P_{\text{macro}} = \frac{1}{M} \sum_{m=1}^M P(TP_m, TN_m, FP_m, FN_m) \quad (2.4.20)$$

Таким чином, для порівняння отриманих результатів навчання з результатами інших авторів для збалансованої вибірки перевагу слід віддавати такій метриці як точність класифікатора, а у випадку незбалансованої вибірки перевагу слід віддавати F-мірі. При цьому для багатокласового класифікатора мікроусереднення метрик продуктивності є більш чутливим до зменшення функціональної ефективності у випадку незбалансованості класів. Для налаштування гіперапараметрів моделі з використання ідей і методів ІЕІ-технології для синтезу вирішувальних правил найбільш доцільним є використання інформаційної логарифмічної міри Шеннона, що може бути виражена як функціонал від значення лічильників результату статистичних тестів класифікатора подібно до інших метрик. При цьому інформаційна міра Шеннона дозволяє прямо оцінювати рівень зняття невизначеності і забезпечує високу узагальнюючу здатність вирішувальних правил за найскладніших у статистичному сенсі умов.

2.5 Моделі екстракції ознакового опису спостережень

Класифікаційний аналіз відеокадрів у найпростішому випадку може бути реалізований у вигляді однієї згорткової мережі, яка може бути навчена як за традиційною схемою, так і у складі сіамських або генеративних моделей.

Для аналізу окремих кадрів пропонується використовувати згорткову мережу загального призначення MobileNet. При цьому коефіцієнт ємності даної мережі встановлено на рівні 0,25 і використовується лише екстрактор ознак (backbone) без

вихідних шарів [86]. Вхідна роздільна здатність мережі може бути обрана із уніфікованого ряду: 160x160 пікселів, 192x192 пікселів, 224x224 пікселів. На рис. 2.10, а показано класичний варіант згорткової мережі, а на рис. 2.10, б модифікований варіант для дослідження пропонованого методу навчання.

Шар Global Average Pooling 2D забезпечує зменшення розмірності, а шар Dropout з коефіцієнтом вимкнення вузлів на рівні 0,5 (тобто 50 %) використовується для регуляризації [87]. Повнозв'язні та сигмоїдні шари утворюють вихідний ознаковий опис.

Однак в ситуаціях втрати видимості чи значного наближення камери до стінок труби отримані зображення втрачають більшу частину корисної контекстної інформації. Тому для відтворення контекстної інформації потрібно кожен кадр аналізувати одночасно з сусідніми кадрами. У загальному випадку модель розпізнавання рівня води вздовж труби матиме у своєму складі на нижньому рівні 2D згорткову мережу для екстракції просторових ознак та 1D темпоральну мережу для аналізу міжкадрових залежностей (рис. 2.11) [14, 16].

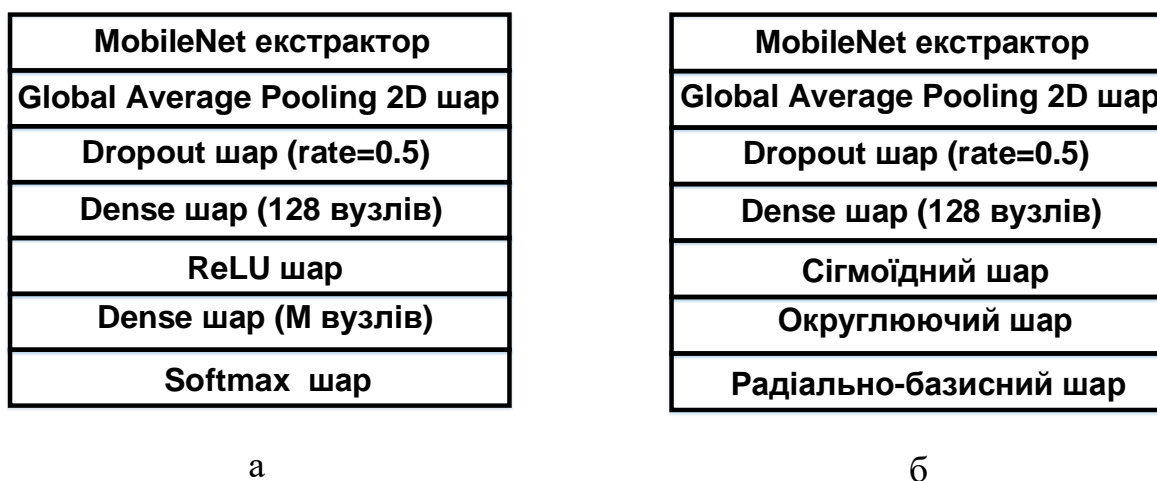


Рисунок 2.10 – Архітектура згорткової мережі :
а – традиційний варіант; б – модифікований варіант

Вирішувальні правила моделі класифікатора зображень містять шар округлення, який формує двійкове подання, та радіально-базисні функції, що визначають належність об'єкта до одного з класів, що розділені гіперсферичними

контейнерами у двійковому просторі Хеммінга.

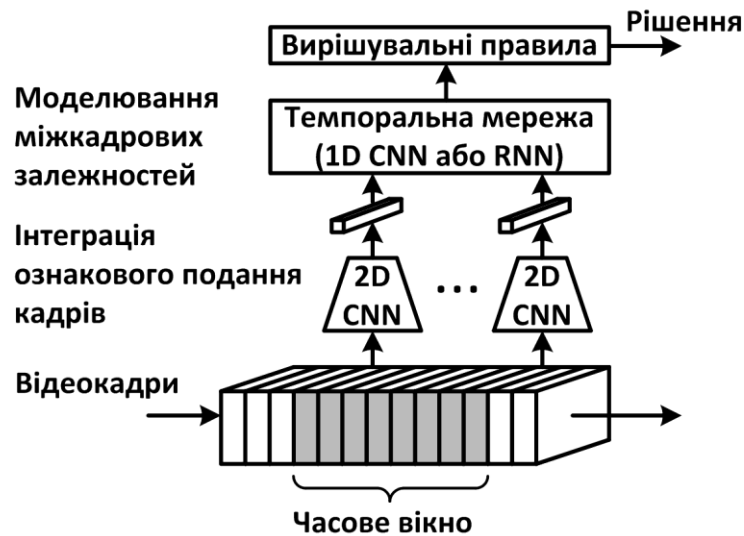


Рисунок 2.11 – Узагальнена архітектура класифікатора рівня води

Кожен гіперсферичний контейнер визначається двійковим еталонним (опорним) вектором, тобто центром контейнера, та радіусом контейнера в кодових одиницях відстані Хеммінга. В цьому випадку радіально-базисна функція належності $\mu_m(\mathbf{b})$ до m -го класу для N -вимірною двійкового вектора \mathbf{b} може бути обчислена за формулою

$$\mu_m(\mathbf{b}) = 1 - \sum_{i=1}^N b_i \oplus b_{m,i}^* / d_m^*, \quad (2.5.1)$$

де \mathbf{b}_m^* – двійковий еталонний вектор (центр оптимального контейнера) для класу X_m^0 ;

d_m^* – радіус оптимального контейнера для класу X_m^0 в кодових одиницях.

Темпоральна мережа може бути реалізована на одній з популярних архітектур, до яких належить 1D регулярна згорткова мережа з дірливими згортками, 1D каузальна згорткова мережа з дірливими згортками, рекурентна мережа з LSTM-комірками, рекурентна мережа з GRU-комірками [81, 88].

Базовий блок (BB) згорткових темпоральних мереж показано на рис. 2.12. Перший шар базового блоку утворюється дірливими 1D згортками (Dilated Conv),

розмір ядра яких дорівнює $k = 3$, а фактор розрідженості (дірявості) подвоюється з кожним наступним базовим блоком, тобто буде дорівнювати 1, 2, 4, 8 і так далі.

На рис. 2.13 показано рецептивне поле каскаду темпоральної мережі з трьома складеними базовими блоками (BB) з регулярними згортками. Регулярні згортки мають рецептивне поле, яке поширюється однаково як праворуч, так і ліворуч. Це означає, що модель дивиться як в майбутнє (праворуч), як і в минуле (ліворуч). Тобто рівень води на поточному кадрі уточнюватиметься на основі інформації як від минулих, так й від майбутніх кадрів.

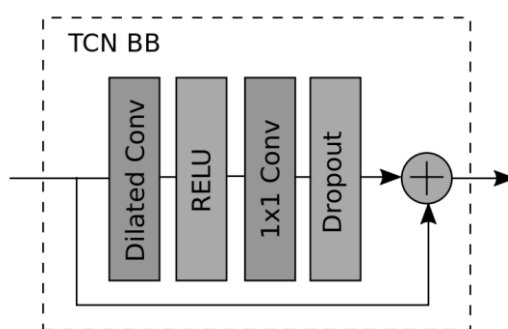


Рисунок 2.12 – Базовий блок темпоральної згорткової мережі

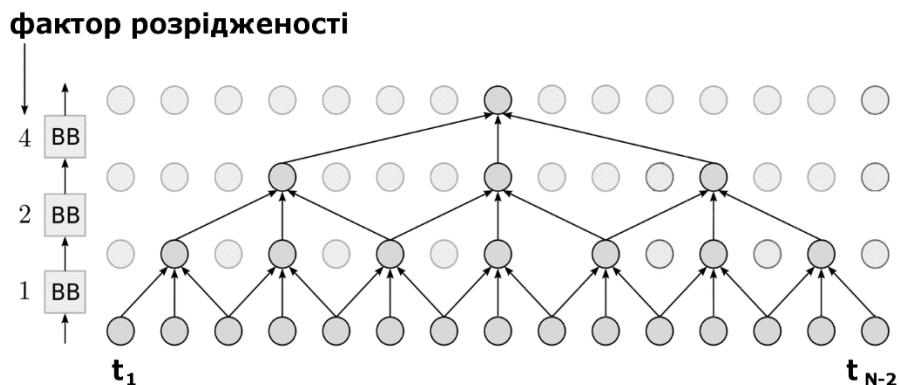


Рисунок 2.13 – Темпоральна мережа з регулярними згортками

Рис. 2.14 ілюструє темпоральну нейромережу з каузальними згортками, що підсилюють продуктивність прогнозування ближче до правого краю. В цьому випадку прогнозований рівень води на поточному кадрі уточнюватиметься на основі ознак і взаємозв'язків попередніх кадрів.

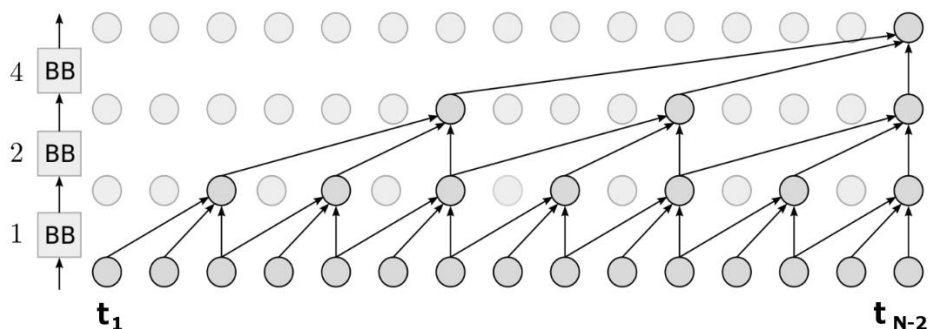
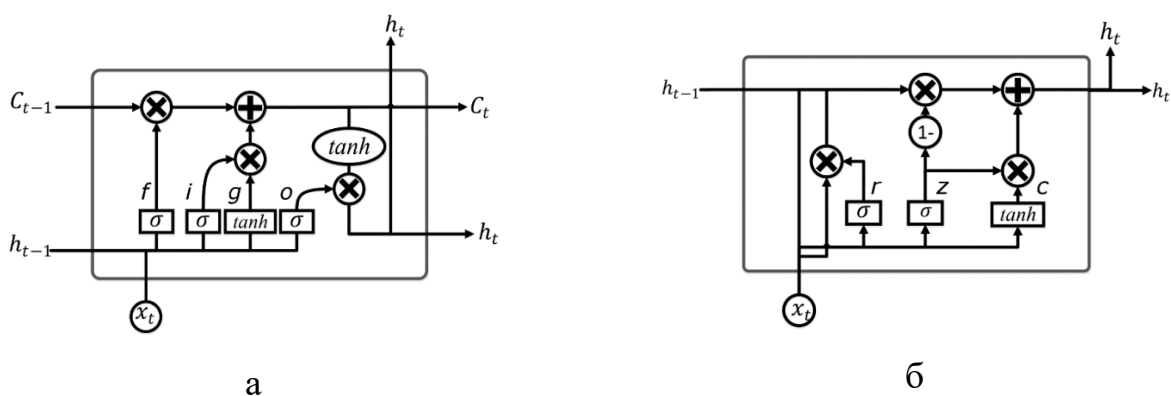


Рисунок 2.14 – Темпоральна мережа з каузальними згортками

LSTM-мережа складається з комірок пам'яті, що з'єднані як послідовні шари, і це дозволяє мережі вибірково запам'ятовувати вхідні навчальні дані через унікальну структуру з трьома вентилями (рис. 2.15, а). Лінія у верхній частині діаграми позначає стан c комірки пам'яті і представляє внутрішню пам'ять. Лінія у нижній частині діаграми – це прихований стан h . Вентилі i , g (вхідні), f (забувальний) та o (вихідний) – це механізми, за допомогою яких LSTM-комірка вирішує проблему зникаючого градієнта. Під час навчання LSTM-комірка оптимізує параметри цих вентилів. Замість вхідних, забувального та вихідного вентилів комірка GRU має лише два вентиля – вентиль оновлення та вентиль скидання (рис. 2.15, б). Вентиль оновлення визначає, скільки попередньої пам'яті слід зберігати, а вентиль скидання визначає, як об'єднати новий вхід із попередньою пам'яттю. В цьому випадку не існує стійкого стану комірки, який би відрізнявся від прихованого, як у комірці LSTM.

Рисунок 2.15 – Комірка рекурентної нейронної мережі :
а – LSTM-комірка; б – GRU-комірка

Моделі, показані на рис. 2.10 та рис. 2.11 можуть навчатися як за традиційним так і за модифікованими методами, що розглядаються в даній дисертаційній роботі. Традиційний метод навчання полягає у додаванні до екстрактора ознак Dense шару з Softmax нормалізацією виходу і використанні методу зворотного поширення помилки, наприклад Adam, з крос-ентропійною функцією втрат. Для моделі з урахуванням просторово-часового контексту (рис. 2.11) спочатку виконується навчання моделі розпізнавання окремого кадру без врахування сусідніх кадрів з використанням традиційного або модифікованого методу навчання. Далі обирається краща навчена модель і використовується її екстрактор ознак, тобто шари, розміщені до сигмоїдного шару включно, для інтеграції кадрів в темпоральну модель розпізнавання. Кожен тип темпоральної нейронної мережі також може навчатися за традиційним або модифікованим методом навчання. При цьому для кожної моделі оптимізується по сітці гіпер-параметр, що відповідає за розмір мережі, тобто за її інформаційну ємність.

Таким чином, запропоновано згорткову модель екстрактора ознак для аналізу як окремих кадрів, так і їх послідовності. При цьому навчання даних моделей можна здійснювати як за традиційними методами навчання так і в рамках нових підходів. Для реалізації завадозахищеного кодування спостережень і побудови інформаційно-екстремальних класифікаційних правил верхній шар нейромереж потребує модифікації шляхом дискретизації ознакового опису та застосування радіально-базисних функцій.

Висновки до другого розділу

У другому розділі проаналізовано особливості процесу відеоінспекції труб водовідведення та питання синтезу моделі інтелектуального аналізу даних для підвищення достовірності звітів про інспекцію. Одержано такі наукові результати:

1. Виявлено необхідність підвищення ефективності існуючих підходів щодо розпізнавання дефектів труб водовідведення. При цьому визначено основні

причини, що ускладнюють завдання інформаційного синтезу ефективної системи аналізу даних, до яких належать :

- ігнорування в існуючих рішеннях контекстної просторово-часової інформації під час синтезу моделей екстракції ознакового опису;

- зашумленість і незбалансованість вибіркового даних та висока варіативність візуальних спостережень в задачі аналізу даних інспекції труб водовідведення;

- рутинність і дороговартісність розмітки достатнього обсягу навчальних даних для ефективного навчання традиційних алгоритмів детектування об'єктів на зображеннях;

- залежність ефективності існуючих методів глибокого машинного навчання від значень гіперпараметрів та початкової ініціалізації вагових коефіцієнтів моделі, а також обсягу і збалансованості навчальних даних;

- неадекватність традиційних критеріїв оптимізації параметрів інтелектуальної системи для складних у статистичному сенсі випадків, обумовлених незбалансованістю та обмеженим обсягом даних;

- непридатність традиційних критеріїв оптимізації параметрів моделі аналізу даних для реалізації інформаційного пляшкового горла з метою зменшення вимог до обсягу розмічених навчальних даних;

- нерентабельність потенційних моделей штучного інтелекту, що забезпечують високу достовірність рішень внаслідок перевищення витрат на навчання моделі та на обробку даних в режимі екзамену над витратами на ручну перевірку відеоінспекцій;

2. Сформовано постановку задачі створення моделей і методів інтелектуальної інформаційної технології для оцінювання функціонального стану труб водовідведення, яка полягає у декомпозиції задачі оцінювання функціонального стану труб на декілька окремих задач класифікаційного аналізу та реалізації ітераційної процедури оптимізації параметрів функціонування за критерієм ефективності. Таким чином, основний напрямок підвищення ефективності моделей оцінювання функціонального стану труб водовідведення полягає в адаптації вхідного математичного опису та побудові в процесі навчання

класифікаційних вирішувальних правил, що призначені для розпізнавання контексту, розпізнавання дефектів в окремих контекстах та оцінювання зміни рівня води, шляхом максимізації критерію ефективності. Для вирішення даної задачі існує потреба забезпечення врахування просторово-часової інформації під час екстракції ознакового опису спостережень та ефективне використання нерозмічених навчальних даних для ініціалізації вагових коефіцієнтів кожної моделі аналізу даних.

3. Розроблено моделі відображення множин та функціональні схеми для опису процесів навчання та розпізнавання функціонального стану труб водовідведення в рамках так званої ІЕІ-технології. У розроблених моделях передбачається багатоетапний процес машинного навчання з вкладеними контурами оптимізації, що забезпечують кодування вхідних спостережень у вигляді інформативного ознакового подання, що зручне для побудови простих і оптимальних в інформаційному сенсі вирішувальних правил. При цьому основними етапами машинного навчання є :

- контрастне навчання екстрактора ознак на нерозмічених навчальних зразках;
- контрастне навчання екстрактора ознак на розмічених навчальних зразках з регуляризацією, що забезпечує мінімізацію помилки дискретизації ознакового опису;
- навчання моделі з заміною міток класів двійковими еталонними векторами;
- формування двійкової навчальної матриці;
- побудова в радіальному базисів двійкового простору ознак інформаційно-екстремального розбиття на класи розпізнавання та обчислення інформаційного критерію для оцінювання ефективності обраних гіперпараметрів моделі.

4. Отримав подальшого розвитку інформаційний критерій ефективності, що обчислюється як функція лічильників результату перевірки статистичних гіпотез, для оптимізації параметрів класифікаційних вирішувальних правил та гіперпараметрів моделі аналізу даних. Запропоновано використовувати мікро- та макро- усереднення значення критерію по множині класів залежно від

збалансованості вибірки. Обґрунтовано використання валідаційних метрик точності та F-міри для порівняння отриманих результатів з результатами інших авторів.

5. Удосконалено моделі екстракції ознакового опису спостережень шляхом поєднання локальної і просторово-часової візуальної контекстної інформації, що дозволяє підвищити інформативність ознакового опису відеоспостережень інспекції труб водовідведення за умов неповної визначеності. Для реалізації заводо захищеного кодування спостережень і побудови інформаційно-екстремальних класифікаційних правил верхній шар нейромереж потребує модифікації шляхом дискретизації ознакового опису та застосування радіально-базисних функцій.

Основні наукові результати, наведені у другому розділі, опубліковано у працях автора [19], [24], [78], [79], [80], [81], [84], [85], [86].

РОЗДІЛ 3

МЕТОДИ ОПТИМІЗАЦІЇ ПАРАМЕТРІВ МОДЕЛІ АНАЛІЗУ ДАНИХ ВІДЕОІНСПЕКЦІЇ ТРУБ ВОДОВІДВЕДЕННЯ

3.1 Метод оптимізації параметрів моделі класифікаційного аналізу зображень з використання нерозмічених та розмічених даних

Оптимізація параметрів екстрактора ознак здійснюється в процесі машинного навчання. Для реалізації моделей процесу машинного навчання, що описані в другому розділі дисертаційного дослідження, пропонується реалізувати алгоритми, в основі яких лежать такі методологічні принципи [89]:

- навчання в декілька послідовних етапів, де кожен попередній етап здійснює ініціалізацію вагових коефіцієнтів моделі для наступного етапу і може повторюватися для реалізації принципу відкладених рішень та узгодження різних методів навчання;

- на першому етапі може здійснюватися навчання без учителя для ефективного використання великого обсягу нерозмічених навчальних даних;

- поєднання дискретного ознакового подання (Deep Hashing) з кодуванням міток класів (Target Coding), тобто один з етапів навчання з учителем повинен сформувати еталонні вектори дискретних ознак, що стануть кодами (мітками) класів для наступного етапу;

- врахування під час навчання екстрактора ознак нелінійності вирішувальних правил, що будуються в радіальному базисі;

- оптимізація параметрів вирішувальних правил повинна здійснюватися на основі інформаційного критерію, оскільки це забезпечує більшу узагальнюючу здатність за умов незбалансованості та обмеженого обсягу розмічених зразків.

Запропонований метод включає 5 основних етапів (фаз) навчання (рис. 3.1), перший з яких є алгоритм навчання без учителя або навчання з самоучителем. В працях [79, 90, 91] розглядаються різні варіанти навчання з самоучителем, найбільш популярним серед яких є методи контрастного навчання і їх модифікації.

У загальному випадку функцією втрат контрастного навчання є NCE (тобто Noise Contrastive Estimation), що забезпечує оцінювання взаємної інформації між ознаковим поданням вибіркового зразків з ознаковим поданням решти зразків.

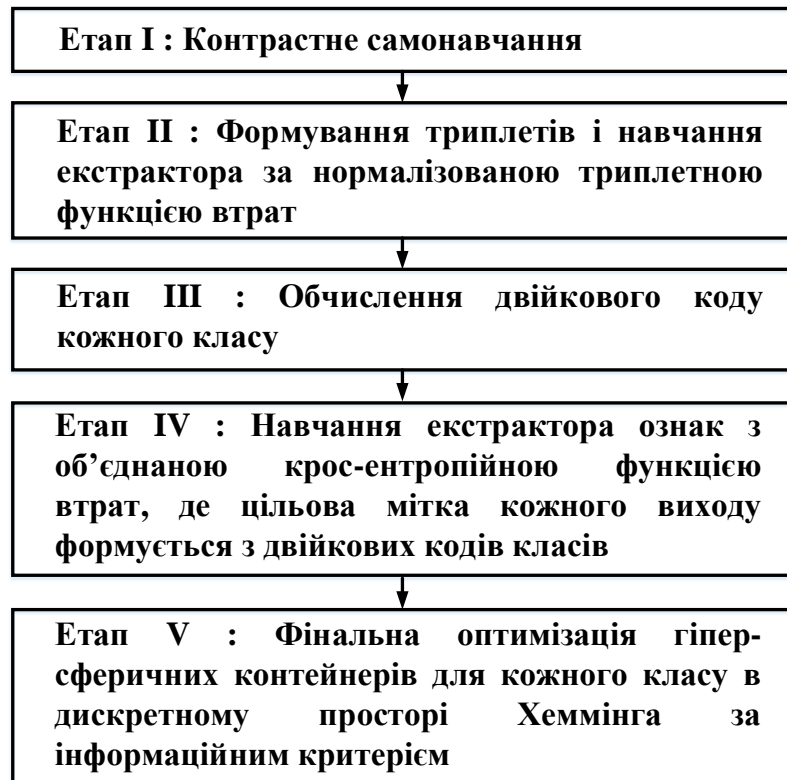


Рисунок 3.1 – Етапи запропонованого методу навчання

Перший етап полягає в навчанні на нерозмічених даних з використанням контрастної функції втрат [90]. Як було показано в дослідженні [91] найбільш біологічно правдоподібна і оптимальна до використання ресурсів пам'яті є екземпляр-прототипна контрастна функція втрат. Алгоритмічна особливість використання контрастної екземпляр-прототипної функції втрат під час навчання (instance-prototype contrastive learning) порівняно з іншими варіантами полягає :

- у використанні неіндексованої FIFO-черги ознакових описів для негативних зразків відносно невеликого обсягу замість індексованих банків пам'яті великого обсягу;
- та у формуванні розширеного набору позитивних пар зразків у вигляді аугментованих версій оригіналів та усереднених за ними прототипів.

Екземпляр-прототипна контрастна функція втрат обчислюється за модифікованою функцією втрат, що має формулу

$$L_{contrastive} = -\frac{1}{n} \sum_{i=1}^n \log \frac{\exp\left(-\frac{1}{N} \|z_i - \bar{z}\|_2^2\right)}{\exp\left(-\frac{1}{N} \|z_i - \bar{z}\|_2^2\right) + \sum_{k=1}^K \exp\left(-\frac{1}{N} \|z_i - z_k\|_2^2\right)}, \quad (3.1.1)$$

де n – кількість зразків в міні-пакеті ($n = 32$);

z_i – ознаковий опис на виході екстрактора ознак $z_i = f(x_i)$ для вхідного x_i зразка з міні-пакету;

$$\bar{z} = \frac{1}{n_a} \sum_{j=1}^{n_a} f(x_j) - \text{усереднене ознакове подання для аугментованих}$$

версій вхідного зразка x_j , що розглядається як позитивна пара для x_j ,

де $f(x)$ – функція, що описує екстрактор ознак і встановлює залежність між вхідним зображенням і вектором на виході сигмоїдного шару;

K – кількість попередньо оброблених зразків, що розглядаються як негативні пари, ознакові подання яких зберігаються в неіндексованому буфері (довжиною $K = 1024$) і оновлюються після обробки кожного міні-пакету;

N – розмірність ознакового опису на виході екстрактора ознак;

$\|\cdot\|_2^2$ – квадрат відстані Евкліда.

Другий етап машинного навчання виконується на розміченому датасеті з використанням триплетної функції втрат у вигляді модифікації NCA функції (Neighbourhood Components Analysis) з майнінгом триплетів. Модифікована NCA функція втрат обчислюється на змішаних міні-пакетах, кожен з яких містить M зображень кожного класу, і має такий вигляд [92]:

$$L_{NCA} = -\log \frac{\exp\left(-\frac{1}{N} \|z_a - z_{ep}\|_2^2\right)}{\exp\left(-\frac{1}{N} \|z_a - z_{ep}\|_2^2\right) + \exp\left(-\frac{1}{N} \|z_a - z_{shn}\|_2^2\right)}, \quad (3.1.2)$$

де $z_a = f(x_a)$, $z_{ep} = f(x_{ep})$ та $z_{shn} = f(x_{shn})$,

де x_a – якірне зображення, обране з міні-паketу;

$$x_{ep} = \arg \min_{x: C(x)=C(x_a)} \|f(x_a) - f(x)\|, \quad (3.1.3)$$

$$x_{shn} = \arg \min_{\substack{C(x) \neq C(x_a) \\ \|f(x_a) - f(x)\| > \|f(x_a) - f(x_p)\|}} \|f(x_a) - f(x)\|, \quad (3.1.4)$$

де $C(x)$ – функція, що повертає мітку класу для конкретного зразка.

Для регуляризації розглядається штрафування за відхилення ознакового подання від двійкового виду. Ця регуляризація забезпечує інформаційне пляшкове горло (Information bottleneck), що сприяє підвищенню завадозахищеності та узагальнюючої здатності моделі. Регуляризуюча складова характеризується Ліпшицевою неперервністю (Lipschitz Continuous) і обчислюється за формулою [6]:

$$R_{bin} = \frac{1}{n} \sum_{i=1}^n \lambda f(x_i)^T (e - f(x_i)), \quad (3.1.5)$$

де e – одиничний вектор;

λ – коефіцієнт регуляризації ($\lambda = 0.01$).

Наступна фаза навчання використовується для обчислення двійкового коду (прототипу) класу для реалізації кодів, що виправляють помилки (error-correcting output codes), але з урахуванням структури класів і відношень між їх ознаками. Для цього навчальна вибірка для M класів $\{x_{m,s} \mid m = \overline{1, M}, s = \overline{1, n_m}\}$, що містить n_m зразків m -го класу, кодується дискретним поданням $\{b_{m,s,i} \mid m = \overline{1, M}, s = \overline{1, n_m}, i = \overline{1, N}\}$ з розмірністю N . Кодування вхідного зображення $x_{m,s}$ двійковим поданням здійснюється операцією округлення i -го виходу сигмоїдного шару [7]

$$b_{m,s,i} = \begin{cases} 1, & \text{if } f_i(x_{m,s}) > 0,5; \\ 0, & \text{otherwise.} \end{cases} \quad (3.1.6)$$

Двійковий еталонний вектор b_m для m -го класу (Label embedding) можна визначити шляхом порозрядного порівняння частоти двійкових одиниць в m -му класі з фоновою частотою одиниць в навчальній вибірці [8, 9]

$$b_{m,i} = \begin{cases} 1, & \text{if } \frac{1}{n_m} \sum_{s=1}^{n_m} b_{m,s,i} > \frac{1}{M} \sum_{c=1}^M \frac{1}{n_c} \sum_{s=1}^{n_c} b_{c,s,i}; \\ 0, & \text{otherwise.} \end{cases} \quad (3.1.7)$$

Еталонний вектор (прототип) z -го класу b_z використовується як цільова розмітка (label) кожного зразка даного класу на наступному етапі навчання з використанням крос-ентропійної функції втрат, яка для вхідного зображення x обчислюється за формулою [10]

$$L = - \sum_{i=1}^N (b_i \log f_i(x) + (1 - b_i) \log(1 - f_i(x))), \quad (3.1.8)$$

де $f_i(x)$ – значення i -го виходу сигмоїдного шару для вхідного зразка x ;

b_i – значення i -го розряду еталонного вектора класу, до якого належить зображення x .

Останній етап машинного навчання призначений для оптимізації кодового радіусу контейнерів за інформаційним критерієм (2.4.17) для врахування дисперсії розподілу кожного з класів в двійковому просторі Хеммніга [85, 22]

$$E_m^* = \max_{\{d\}} E_m(d), \quad (3.1.9)$$

де $\{d\} = \{0, 1, \dots, \left(\sum_i b_{m,i} \oplus b_{c,i} - 1 \right)\}$ – набір радіусів з центром в b_m ;

E_m – інформаційний критерій для m -го класу, який є функцією від характеристик точності [22].

Параметри функціонування системи аналізу даних можна поділити на чотири основні групи:

- гіперпараметри, що впливають на ємність моделі аналізу даних;
- гіперпараметри, що впливають на поведінку алгоритму навчання;
- параметри збору, зберігання і попередньої обробки даних;

До гіперпараметрів, що впливають на процес попередньої обробки, можуть належати частота опитування сенсорів, роздільна здатність сенсорних систем, параметри аугментації даних та інші. До гіперпараметрів, що впливають на ємність моделі, належать кількість нейронів, кількість шарів, розмір композиції моделей та інше. До параметрів, що впливають на поведінку алгоритму, належать швидкість навчання, коефіцієнт зниження швидкості навчання, параметри регуляризації та інші [79, 85, 22]. Під час оптимізації гіперпараметрів оцінку їх ефективності здійснюють на зовнішній валідаційній вибірці, що не брала участі у навчанні.

Найбільш простим методом оптимізації параметрів функціонування системи аналізу даних є метод пошуку по сітці (Grid Search), який виконує повний перебір по заданій вручну підмножині простору параметрів [95]. Пошук по сітці повинен супроводжуватися деяким вимірюванням продуктивності за комплексним критерієм або за точнісною характеристикою, обчисленою на валідаційних вибірках. Для прискорення пошуку також інколи використовують метод випадкового пошуку (Random Search). В рамках даного методу повний перебір всіх комбінацій замінюється на їх вибірку випадковим чином. Даний метод може застосовуватися як для дискретної множини значень, так і для неперервних чи змішаних просторів [96].

Також існує практика використання методу Байєсовської оптимізації (Bayesian Optimization), що полягає в побудові стохастичної моделі функції відображення значень (гіпер)параметра в цільову функцію [97]. Алгоритм має ітеративну структуру, на кожній ітерації якої перспективна конфігурація параметрів моделі тестується і оновлюється. При цьому в алгоритмі передбачено балансування між зондуванням простору рішень і використанням зібраної інформації про близькість до оптимуму. Даний метод перевершує результати пошуку по сітці та випадкового пошуку як з точки зору кількості обчислень, так і з точки зору оптимальності знайденого рішення.

Таким чином, процес машинного навчання моделі класифікаційного аналізу даних полягає в ініціалізації вагових коефіцієнтів за принципами переносу знань та контрастного навчання без вчителя, уточненні вагових коефіцієнтів на розміченій

вибірці з одночасним визначенням кодового подання кожного з класів (label embedding) та фінальному уточненні вагових коефіцієнтів з використанням отриманого кодування цільової розмітки. При цьому останній етап машинного навчання полягає в оптимізації кодових радіусів кожного класу розпізнавання для уточнення меж класів у просторі ознак. Вибір гіперпараметрів може бути здійснений за традиційними методами пошуку по сітці, випадкового пошуку чи шляхом Байєсівської оптимізації.

3.2 Оптимізація параметрів моделі розпізнавання контекстів спостереження

Для врахування орієнтації камери та функціонального стану системи відеомоніторингу під час кодування звітів про інспекцію труб доцільно використовувати класифікатор контексту спостережень, що детально описано в підрозділі 2.1. Алфавіт класів, що використовується для визначення контексту спостереження без урахування вмісту сусідніх кадрів, містить 10 основних класів контексту (табл. 3.1) [80].

Таблиця 3.1 – Множина класів для класифікатора контексту окремих кадрів

Позначення класу	Кількість зразків	Назва контексту	Опис
X_1^o	5000	Forward	Камера спрямована вперед уздовж труби
X_2^o	3000	Side	Камера спрямована на стінку труби, коли важко зрозуміти, на яку частину труби (зверху, знизу, зліва чи справа) спрямована камера.
X_3^o	1000	Semi right	Неповний правий поворот, при якому орієнтація камери чітко визначається з одного кадру.

Продовження таблиці 3.1

Позначення класу	Кількість зразків	Назва контексту	Опис
X_4^o	1000	Semi left	Неповний лівий поворот, при якому орієнтація камери чітко визначається з одного кадру.
X_5^o	320	Semi top	Неповний поворот вгору, при якому орієнтація камери чітко визначається з одного кадру.
X_6^o	180	Semi bottom	Неповний поворот вниз, орієнтація камери чітко визначається з одного кадру.
X_7^o	2100	Connection	Інша труба з'єднується з основною, як правило, це труба меншого діаметру, що прямує від домогосподарств
X_8^o	500	Manhole	Місце входу в трубу для огляду, від оглядового колодязя до початку самої труби
X_9^o	152	Collapse	Обвалилася труба, подальший рух вперед неможливий
X_{10}^o	100	Ignore	Ситуації, які слід ігнорувати для цілей обробки

Сформований алфавіт містить клас ігнорування, клас вид збоку та під'єднання. Вибірki для цих класів легко збирати та маркувати, проте вони не дають повної впевненості щодо орієнтації камери, оскільки не зрозуміло, до якої стінки труби звернена камера. Тому алфавіт класів, що використовується для визначення контексту спостереження з урахуванням вмісту сусідніх кадрів, містить 11 основних класів контексту (рис. 2.3). Для проведення експериментального дослідження було використано дані, отримані від компанії Ace Pipe Cleaning, Inc. Кількість розмічених зразків кожного класу приведено в табл. 3.2 [80].

Таблиця 3.2 – Множина класів для класифікатора контексту з урахуванням сусідніх кадрів

Позна-чення класу	Кількість зразків	Назва контексту	Опис
X_1^o	4500	Forward	Камера спрямована вперед вздовж труби
X_2^o	1500	Right	Неповний або повний поворот камери вправо
X_3^o	1420	Left	Неповний або повний поворот камери вліво
X_4^o	400	Up	Неповний або повний поворот камери вгору
X_5^o	180	Down	Неповний або повний поворот камери до низу
X_6^o	500	Right connection	Під'єднання до труби справа
X_7^o	500	Left Connection	Під'єднання до труби зліва
X_8^o	150	Top connection	Під'єднання до труби вгорі
X_9^o	75	Down connection	Під'єднання до труби до низу
X_{10}^o	500	Manhole	Місце входу в трубу для огляду, від оглядового колодязя до початку самої труби
X_{11}^o	170	Collapse	Обвалилася труба, подальший рух вперед неможливий

Перед навчанням весь набір даних балансується шляхом застосування аугментації до міноритарних за обсягом вибірки класів (0-5% зміна масштабу, $\pm 5\%$ поворот, $\pm 5\%$ зміна яскравості). Кожна модель має ряд гіперпараметрів, які визначають її конфігурацію та потужність. Оптимальні за усередненим значенням інформаційного критерію (2.4.16) результати отримані на роздільній здатності 160x160 пікселів і саме ці результати будуть проілюстровані.

Спочатку навчається модель класифікації контексту окремого кадру, яка потім стає складовою класифікатора, що уточнює результат за рахунок врахування сусідніх кадрів. Вибір архітектури темпоральної мережі, що моделює міжкадрові залежності, відбувається за результатами навчання з оптимально підібраними гіперпараметрами. Моделі навчаються протягом 60 епох. Гіперпараметри підбираються таким чином, щоб максимізувати інформаційний критерій і уникнути ефекту перенавчання.

На рис. 3.2 показано зміну мікроусередненого (2.4.19) значення F-міри (2.4.7) на тестовому наборі даних під час навчання класифікатора за базовим (baseline) та запропонованим (proposed) методами. Базовий алгоритм полягає в одноетапному навчанні традиційної архітектури (2.10а), а запропонований метод полягає у використанні архітектури (2.10б) та багатоетапного алгоритму, що описаний блок-схемою на рис.3.1. Тестовий набір даних було створено шляхом відбору 15% вибірок з кожного класу, представленого в табл. 3.1 [80].

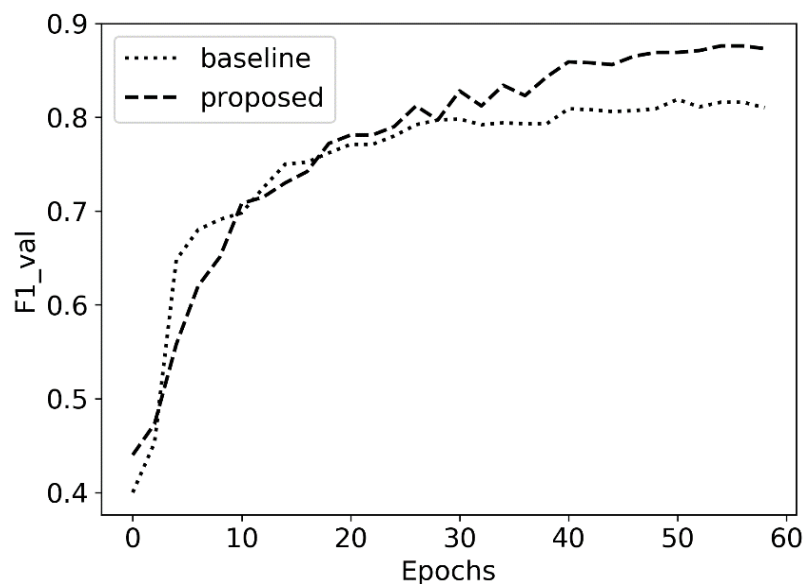


Рисунок 3.2 – Графік залежності мікроусередненого значення F-міри (2.4.7) на тестовому наборі даних від кількості епох навчання для класифікатора контексту кадру без урахування сусідніх кадрів

Аналіз рис. 3.2 показує, що до 30 епохи обидва методи навчання працюють з однаковою ефективністю. Однак третій етап (рис. 3.2) запропонованого методу, що

застосовується після 30-ї епохи, підвищує значення F1-міри на 6%, яка досягає значення 0,878. Тобто використання екстрактора ознак, навченого за запропонованим методом, та обчислення кодових міток класів дозволяє суттєво покращити точність отриманої моделі класифікатора.

Для підвищення інформативності ознакового опису пропонується враховувати темпоральні ознаки шляхом агрегації ознакового опису послідовних (сусідніх) кадрів у певному часовому вікні (рис. 2.11). Часове вікно для розгляду сусідніх кадрів становить 128 кадрів. Це вікно було обрано кратним 2 і підібрано експериментально як близьке до оптимального значення для різних моделей.

Темпоральна згорткова мережа має 1 ступінчасту конфігурацію з 7-ма базовими блоками (рис. 2.12). Кількість каналів ознак у згорткових фільтрах першого шару задано як $C=128$, що відповідає розмірності ознакового подання (вбудовування) зображення. Останній шар темпоральної мережі підключено до повнозв'язного (Dense) шару, який містить 128 вузлів з сигмоїдальною функцією активації.

Рекурентні мережі GRU та LSTM містять по одному шару з 128 вузлами. Стани кожного блоку подаються на шар Dense, який також містить 128 вузлів з сигмоїдною функцією активації. Після сигмоїдного шару будь-якої з моделей використовуються два додаткові вихідні шари, аналогічні зображеним на рис. 2.10, в залежності від обраного методу навчання.

На рис. 3.3 показано зміну мікроусередненого значення F-міри (2.4.7) на тестовому наборі даних під час навчання класифікатора кадрів, який враховує сусідні кадри [80]. При цьому на графіку показано криві, які відповідають згортковій регулярній мережі (F1_val_cnn_regular), згортковій каузальній мережі (F1_val_cnn_regular), рекурентній GRU-мережі (F1_val_gru) та рекурентній LSTM-мережі (F1_val_lstm) відповідно. Навчання проводилося за базовим методом, але з різною архітектурою темпоральної мережі. Тестовий набір даних створюється аналогічно, але з відбором 15% вибірок кожного класу.

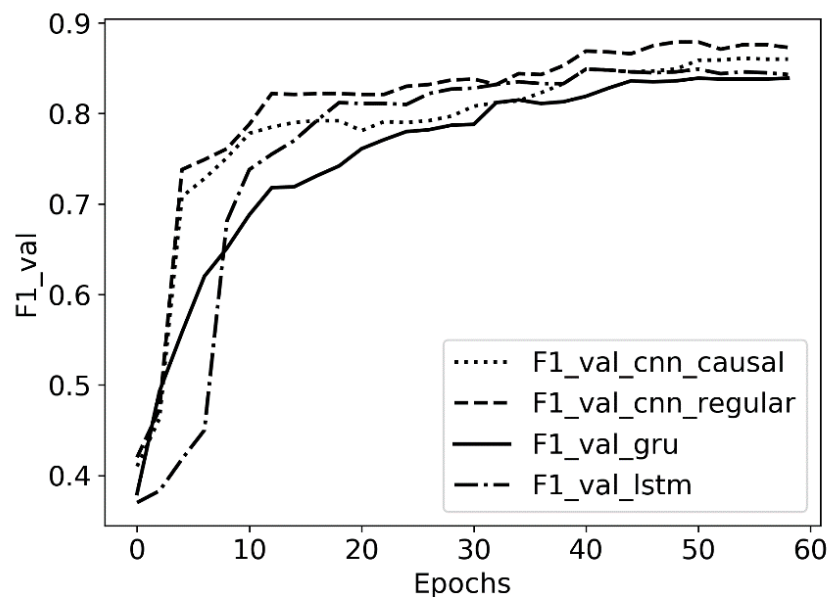


Рисунок 3.3 – Графік залежності мікроусередненого значення F-міри (2.4.7) на тестовому наборі даних від кількості епох навчання за традиційним алгоритмом для класифікатора контексту з урахування сусідніх кадрів

Аналіз рис. 3.3 показує, що точність рекурентних мереж на початкових епохах навчання зростала повільніше, ніж точність згорткових мереж. Однак на завершальних етапах точність фактично вийшла на плато, причому результати для різних рекурентних мереж наближаються до практично однакових. Серед згорткових мереж найкращий результат показала регулярна (рис. 2.13), а не каузальна архітектура (рис. 2.14). Мікроусереднене значення F-міри для регулярної темпоральної згорткової мережі досягло 0,87, тоді як мікроусереднене значення F-міри для LSTM та GRU моделі становить 0,843 та 0,839 відповідно.

На рис. 3.4 показано зміну мікроусередненого значення F-міри (2.4.7) на тестовому наборі даних під час навчання класифікатора, що враховує сусідні кадри, за запропонованим методом (рис. 3.1) з використанням різних архітектур темпоральних мереж [80].

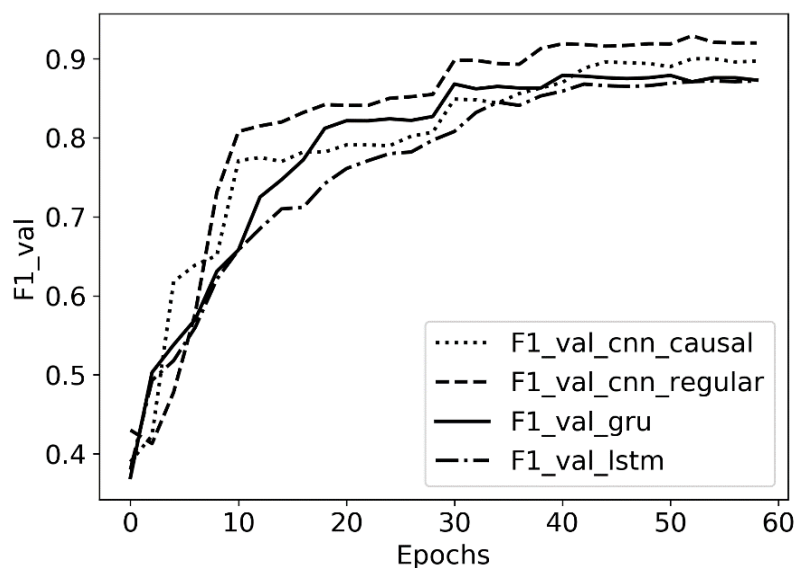


Рисунок 3.4 – Графік залежності мікроусередненого значення F-міри (2.4.7) на тестовому наборі даних від кількості епох навчання за запропонованим алгоритмом (рис. 3.1) для класифікатора контексту з урахування сусідніх кадрів

Аналіз рис. 3.4 показує, що точність рекурентних мереж у даному випадку також поступається точності згорткових мереж. Модель з регулярною структурою (рис. 2.13) також виявилася лідером серед згорткових мереж. Однак максимальне мікроусереднене значення F-міри для моделі з регулярною структурою, навченої запропонованим методом, становить 0,92, що на 4% перевищує результати навчання базового методу на тій самій моделі мережі.

Для адекватної оцінки ефективності запропонованого підходу варто розглядати не тільки агреговані метрики, але й матрицю помилок. Нормалізована матриця помилок, отримана на моделі з регулярними згортковими темпоральними мережами, навченими за традиційним підходом, наведена на рис. 3.5 [80].

Аналіз матриці помилок, представленої на рис. 3.5, показує, що найнижча точність відповідає складним класам: під'єднання справа, під'єднання зліва, під'єднання зверху та колодязь. Ці класи поєднують в собі як орієнтацію камери, так і характерні конструктивні елементи каналізаційної системи. При цьому

мікросереднє значення точності (2.4.6) моделі становить 88,1%, а мікросереднє значення інформаційного критерію (2.4.17) становить 0,57.

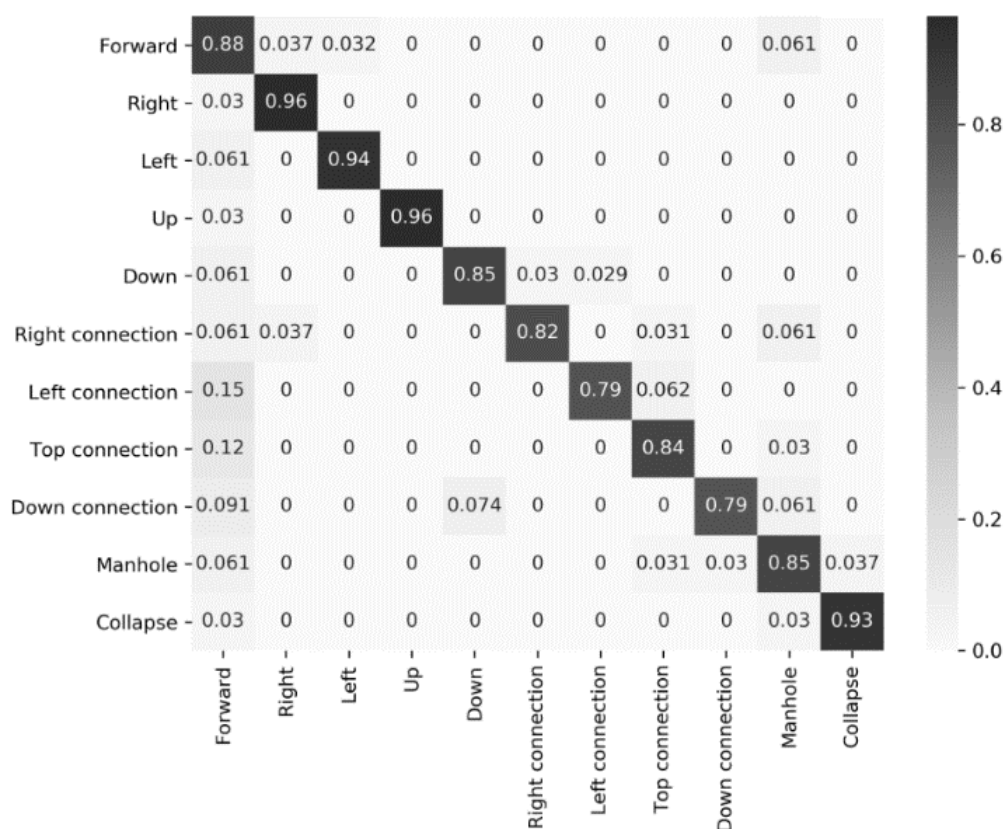


Рисунок 3.5 – Нормалізована матриця помилок для оптимального класифікатора контексту, побудованого з використанням регулярних згорткових темпоральних мереж, навчених за традиційним підходом

На рис. 3.6 зображено нормалізовану матрицю помилок для класифікатора контексту на основі, побудованого з використанням регулярних згорткових темпоральних мереж, що навчений за запропонованим підходом (рис. 3.1) [80].

Аналіз рис. 3.6 показує, що клас Forward перетинається майже з усіма іншими класами. У порівнянні з традиційним методом навчання, запропонований підхід забезпечує підвищення точності розпізнавання для складних контекстних класів, таких як Right Connection, Left Connection, Top Connection та Manhole. При цьому мікросереднє значення точності (2.4.6) моделі становить 0,93 а мікросереднє значення інформаційного критерію (2.4.17) становить 0,67.

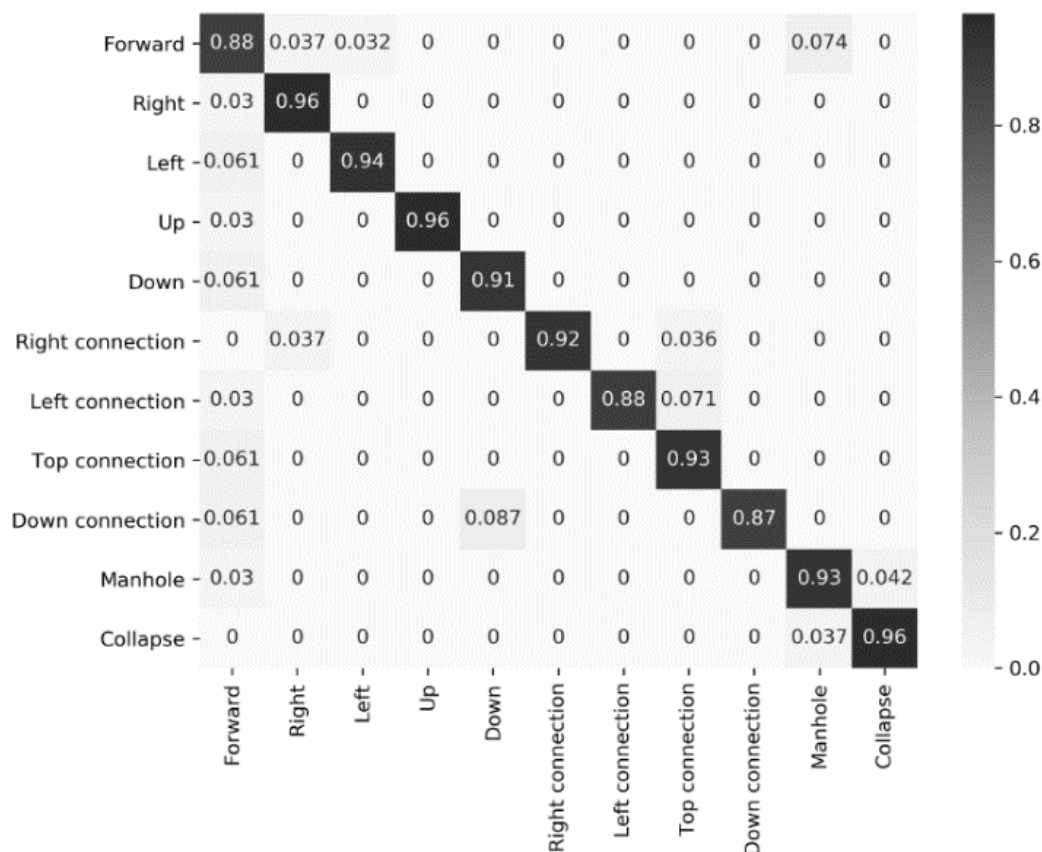


Рисунок 3.6 – Нормалізована матриця помилок для оптимального класифікатора контексту, побудованого з використанням регулярних згорткових темпоральних мереж, навчених за запропонованим підходом

Незважаючи на підвищення точності, навчання на одних і тих же даних все ще не є безпомилковим. Деякі з'єднання (зразки класу Connection) можуть бути пропущені, а для з'єднань в секторі 10:00-02:00 за годинниковою стрілкою орієнтація камери може бути переплутана зі з'єднаннями в секторах 02:00-05:00 та 07:00-10:00 за годинниковою стрілкою. Це може бути пов'язано з незначною різницею між конкретними значеннями кута часткового повороту камери. Однак вплив помилкової класифікації контексту на перетині класів контексту може нівелюватися за рахунок агрегації результатів розпізнавання з однаковим значенням одометра.

Таким чином, встановлено перевагу використання згорткової моделі для аналізу контексту спостереження та запропонованого методу багатоетапного навчання. Також емпірично доведено, що регулярна згорткова архітектура

темперальної мережі є найбільш ефективною при використанні міжкадрових залежностей. Отримана точність є придатною для практичного використання, оскільки додаткова корекція похибки може бути здійснена за рахунок використання даних одометра/лічильника відстані.

3.3 Оптимізація параметрів моделі оцінки рівня води в трубопроводі водовідведення

Оцінка рівня води в трубопроводах водовідведення є важливою складовою інспекції. Рівень води є ключовим показником доступності системи водовідведення для проведення оцінки функціонального стану. Зависокий рівень води приховує дефекти труби і знижує ефективність інспекції. Крім того, визначення зміни рівня води вздовж труби під час інспекції дозволяє виявляти місця просадки або порушення цілісності труби. Просадка труби є одним з важливих кодів функціонального стану труби в рамках американського та британського стандартів [79]. На перший погляд, оцінювання рівня води в трубах водовідведення може здатися простим завданням для алгоритмів машинного зору, оскільки труби мають досить просту конструкцію з невеликою кількістю під'єднань. Однак відеозображення містять значну шумову складову через туман, випаровування, відблиски світла та втрату видимості під водою. Крім того, камера спостереження під час інспекції здійснює повороти для детального огляду зон інтересу. Все це обумовлює необхідність врахування під час аналізу даних зміни орієнтації камери та видимості на послідовності кадрів інспекції.

Алфавіт класів для розпізнавання рівня води в трубі на кадрі зображення без врахування зображень сусідніх кадрів містить 12 класів, що відповідають рівням заповнення труби водою у відсотках: 0 %, 5 %, 10 %, 15 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 % (рис. 3.7) [81].

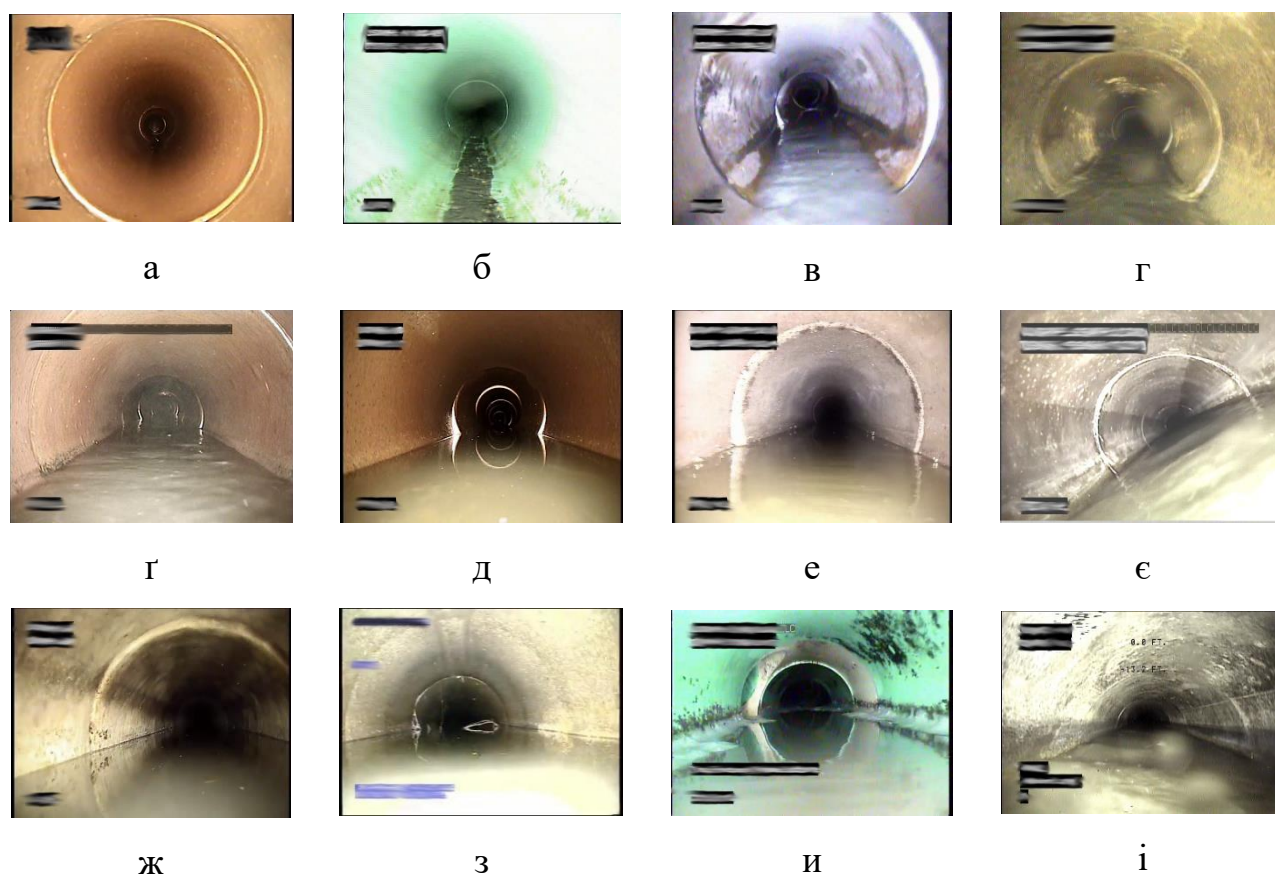


Рисунок 3.7. – Приклади зображень кожного з класів: а – клас X_1^0 ;

б – клас X_2^0 ; в – клас X_3^0 ; г – клас X_4^0 ; г – клас X_5^0 ; д – клас X_6^0 ;

е – клас X_7^0 ; е – клас X_8^0 ; ж – клас X_9^0 ; з – клас X_{10}^0 ;

и – клас X_{11}^0 ; i – клас X_{12}^0

Парсинг розмічених відеофайлів $\{D_v^{train}\}$ та $\{D_w^{test}\}$ призводить до формування різної кількості зразків для кожного класу. Перед навчанням весь датасет балансується шляхом застосування аугментації міноритарних класів. Як методи аугментації використовуються: зміна масштабу на 1...5 %, поворот зображення на $\pm 18^\circ$, зміна яскравості на $\pm 5\%$. Розмір аугментованої навчальної вибірки кожного класу для класифікатора кадрів становить 3000 зразків.

Архітектура моделей обирається аналогічною до моделей, використовуваних в підрозділі 3.2. Кожна модель має ряд гіперпараметрів, що визначають конфігурацію і ємність моделі. Спочатку здійснюється підбір оптимальних гіперпараметрів для кожної архітектури, а потім порівняльний аналіз результатів, отриманих на різних архітектурах з цими параметрами. Навчання моделей

відбувається протягом 60 епох. При цьому гіперпараметри підібрані так, щоб уникнути помітного ефекту перенавчання.

Рис. 3.8 показує зміну мікроусередненого значення F-міри (2.4.7) на тестовій вибірці під час навчання класифікатора окремих кадрів за традиційним (baseline) та запропонованим (рис. 3.1) методами [81]. При цьому тестова вибірка формувалась шляхом вибірки з набору кожного класу 15 % оригінальних зразків.

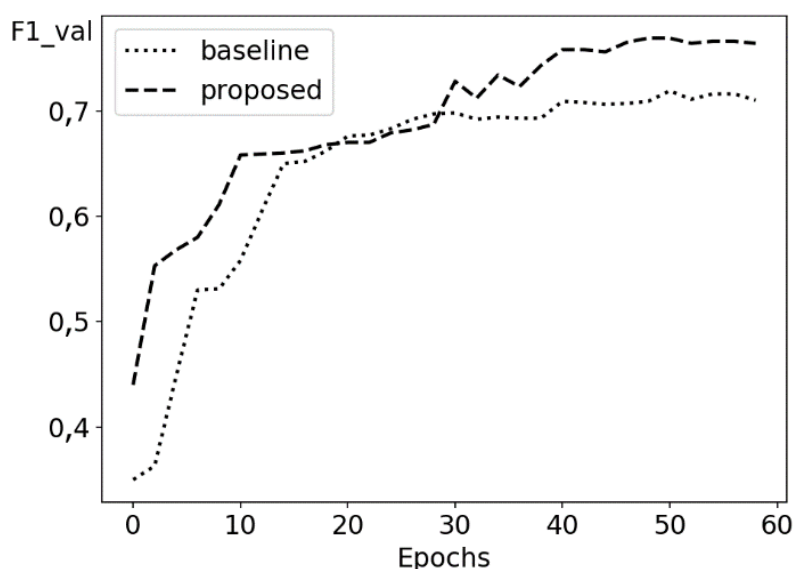


Рисунок 3.8 – Графік залежності мікроусередненого значення F-міри (2.4.7) на тестовій вибірці від кількості епох навчання для класифікатора контексту без врахування сусідніх кадрів

Аналіз рис. 3.8 показує, що до 30-ї епохи ефективність навчання обох методів є приблизно однаковою. Однак після 30-ї епохи навчання у запропонованому методі використовується 3-тя фаза (рис. 3.1), яка забезпечує підвищення мікроусередненого значення F-міри (2.4.7) моделі на 5,4 %, яка становить 0,781. Тому для інтеграції ознакового опису кадрів у високорівневу нейромережу буде використано екстрактор ознак, навчений в рамках запропонованого методу.

Алфавіт класів для розпізнавання рівня води з урахуванням сусідніх кадрів містить ті самі класи. Часове вікно для врахування сусідства дорівнює 257 (тобто $128 \cdot 2 + 1$) кадрам. Розмір часового вікна обирається як число, кратне 2, з додаванням одного кадру, щоб забезпечити симетричність для регулярної

структури моделі. При цьому розмір вікна підбирається експериментально, щоб бути близьким до оптимального для всіх моделей.

Рис. 3.9 ілюструє зміни мікроусередненого значення F-міри (2.4.7) на тестовій вибірці ($F1_{val}$) під час навчання класифікатора кадрів з урахуванням сусідніх кадрів за традиційним (baseline) методом, але для різних архітектур темпоральної мережі: темпоральної мережі з каузальними згортками ($F1_{val_cnn_causal}$); темпоральної мережі з регулярними згортками ($F1_{val_cnn_regular}$); рекурентної мережі GRU ($F1_{val_gru}$); рекурентної мережі LSTM ($F1_{val_lstm}$) [81]. Тестова вибірка формувалась так само шляхом вибірки з набору кожного класу 15 % зразків. Аугментована навчальна вибірка кожного класу становить 2000 зразків.

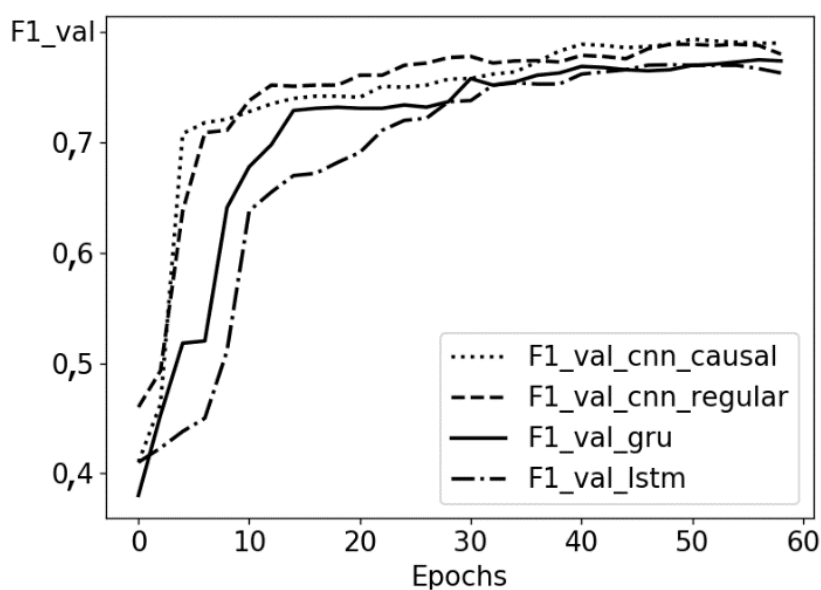


Рисунок 3.9 – Графік залежності мікроусередненого значення F-міри (2.4.7) на тестовій вибірці від кількості епох навчання за традиційним методом навчання для класифікатора контексту з урахуванням сусідніх кадрів

Аналіз рис. 3.9 показує, що точнісні характеристики рекурентних мереж на перших епохах навчання зростали повільніше ніж у згорткових. На останніх епохах навчання зростання ефективності рекурентних мереж припинилося і досягло практично однакових значень. Згорткові мережі також припинили нарощування

ефективності на останніх епохах, однак забезпечили дещо вищі мікроусереднені значення F-міри (2.4.7), але суттєвої різниці між ефективністю каузальної та регулярна архітектури не спостерігається. Отримане мікроусереднене значення F-міри (2.4.7) для регулярної темпоральної мережі становить 0,789, в той час як мікроусереднене значення F-міри (2.4.7) для LSTM та GRU моделей становить 0,770 та 0,775 відповідно. При цьому для регулярної темпоральної мережі мікроусереднене значення точності (2.4.6) моделі становить 0,82, а мікроусереднене значення інформаційного критерію (2.4.17) становить 0,31.

Рис. 3.10 показує зміну мікроусередненого значення F-міри (2.4.7) на тестовій вибірці під час навчання класифікатора кадрів з урахуванням сусідніх кадрів за запропонованим методом (рис. 3.1), але для різних архітектур темпоральної мережі (таких же, як на рис. 3.9) [81].

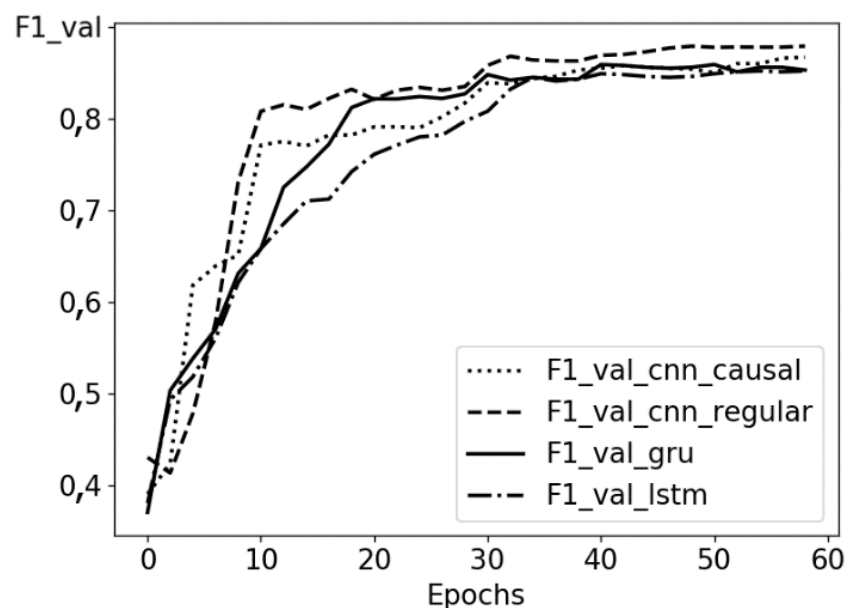


Рисунок 3.10 – Графік залежності мікроусередненого значення F-міри (2.4.7) на тестовій вибірці від кількості епох навчання за запропонованим методом для класифікатора контексту з урахуванням сусідніх кадрів

Аналіз рис. 3.10 показує, що точнісні характеристики рекурентних мереж так само поступаються точнісним характеристикам згорткових мереж. Серед згорткових мереж лідером є модель з регулярною структурою. Максимальне

мікроусереднене значення F-міри (2.4.7) для моделі з регулярною структурою за запропонованим методом навчання дорівнює 0,879, що на 9 % перевищує результат традиційного методу навчання. При цьому мікроусереднене значення точності (2.4.6) моделі становить 0,88, а мікроусереднене значення інформаційного критерію (2.4.17) становить 0,47.

Для останньої моделі з регулярною згортковою темпоральною мережею, навченою в рамках запропонованого підходу, також було обчислено макроусереднене (2.4.20) значення F-міри (2.4.7) для класів таких рівнів води: 0 %, 10 %, 20 %, 30 %, 40 %, 50 %. Це значення становить 0,812. Його можна порівняти з обчисленим макроусередненим (2.4.20) значенням F-міри (2.4.7) для тих самих класів, що приведено у праці [71] і становить 0,36. У праці [98] було запропоновано інший метод побудови класифікатора рівня води, але отримані результати поступаються навіть результатам отриманим в праці [71].

На рис. 3.11 показано матрицю помилок (confusion matrix) для отриманого класифікатора контексту з урахуванням сусідніх кадрів на основі темпоральної згорткової мережі з регулярною структурою згорток.

Аналіз рис. 3.11 показує, що в основному перетинаються сусідні класи, тобто помилка виникає на межі переходу одного рівня води до іншого. Цей результат є очікуваний і може бути пов'язаний з помилками розмітки даних внаслідок суб'єктивного фактору оператора. Також на зображеннях труб з чистою прозорою водою є хибне розпізнавання рівня 0 %.

Набори даних можуть бути незбалансованими щодо форми труби, матеріалу труби та стилю текстових надписів (анотацій), що формуються анотаторами системи збору даних. Ці фактори і низка інших можуть стати причиною появи хибних кореляцій в ознаковому описі. Наприклад, алгоритм може вирішити, що цегла як матеріал труби може бути ознакою 10% рівня заповнення води, якщо у вибірці даних з цим матеріалом більшість зображень мають рівень води 10%.

Як було описано в першому розділі дисертаційної роботи під час формування чи попередньої підготовки ознакового опису є два основні підходи – формування інформативного ознакового подання та формування інваріантного ознакового

подання. Хоча в конкретних прикладних задачах досягнути обидві мети можна одночасно.

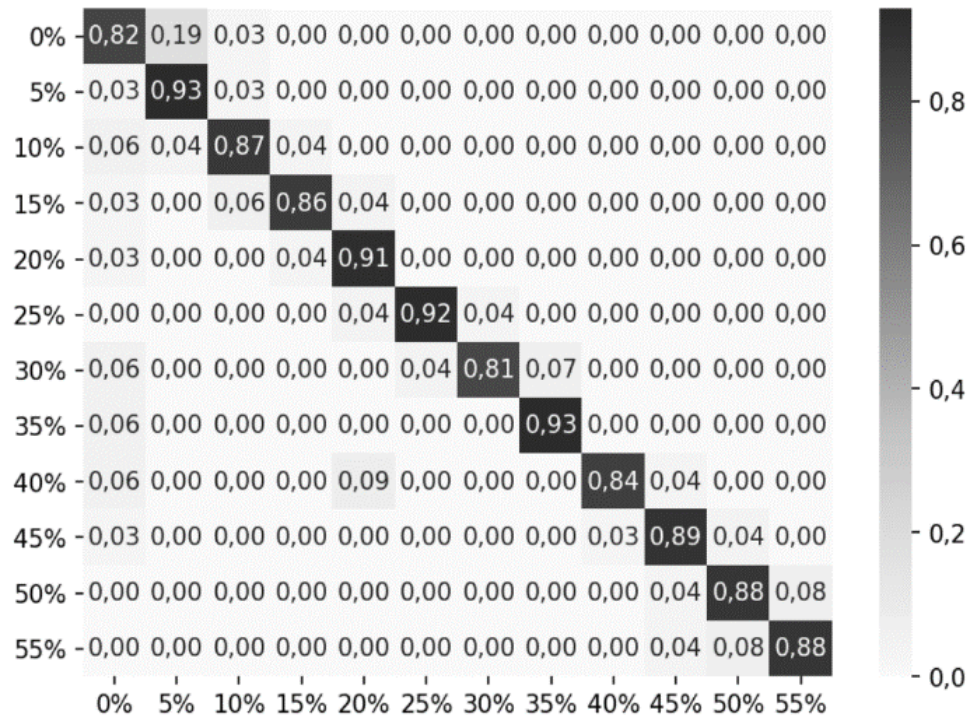


Рисунок 3.11 – Матриця помилок для оптимального класифікатора рівнів води з урахуванням сусідніх кадрів

Результати, наведені на рис. 3.8-3.10, отримані з використанням попереднього самонавчання екстрактора ознак на основі контрастно-прототипної функції втрат (3.1.1). Даний підхід спрямовано на формування інформативного ознакового подання, яке у загальному випадку може бути надлишковим. Для усунення хибної кореляції в даних, також варто розглянути такий метод попереднього навчання, що потенційно може підвищити інваріантність до даних факторів. Пропонується розглянути реалізацію попереднього навчання екстрактора ознак в складі сегментатора води в трубі, щоб екстрактор ознак максимально фокусувався на воді. Як архітектуру сегментатора пропонується використати U-net [99]. На рис. 3.12 показано приклад зображень і їх розмітки для навчання сегментатора води. У табл. 3.3 приведено результати навчання за запропонованим алгоритмом для різних варіантів ініціалізації вагових коефіцієнтів мережі [100].

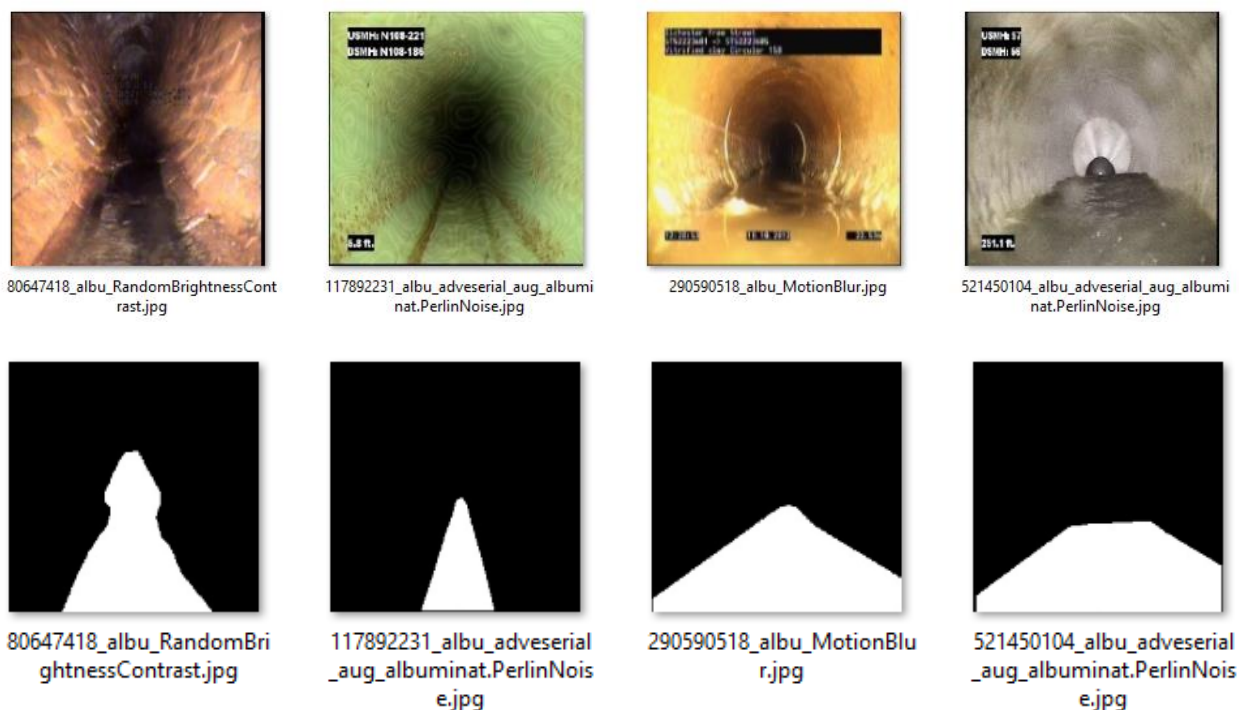


Рисунок 3.12 – Приклади розмічених зображень для навчання сегментатора води в трубопроводі водовідведення

Таблиця 3.3 – Залежність метрики F1 на тестовій вибірці від способу ініціалізації екстрактора ознак

Спосіб ініціалізації	Мікроусереднене значення F-міри (2.4.7) на тестовому наборі даних
Випадкові ваги	0,819
Попереднє навчання на наборі даних ImageNet	0,822
Попереднє навчання в складі сегментатора води	0,861
Самонавчання на основі контрастно-прототипної функції втрат	0,879

Аналіз табл. 3.3 показує, що ініціалізація вагових коефіцієнтів в процесі самонавчання з екземпляр-прототипною контрастною функцією втрат (3.1.1) забезпечує найкращі результати [100]. Це може пояснюватися кращим

узагальненням даних порівняно з іншими підходами. Ініціалізація екстрактора ознак в складі сегментатора ймовірно призводить до втрати частини важливої інформації про форму труби, кут огляду камери.

Таким чином, доведено перевагу використання запропонованої моделі для аналізу рівня води і запропонованого багатоетапного методу її навчання. Експериментально показано, що найбільш ефективною моделлю для класифікаційного аналізу рівня води на відеокадрах є комбінація згорткової 2D мережі і темпоральної згорткової мережі з регулярними згортками. При цьому згорткова темпоральна мережа продемонструвала вищу ефективність врахування міжкадрових залежностей порівняно з рекурентними мережами. Також продемонстровано перевагу у використанні екземпляр-прототипної контрастної функції втрат для ініціалізації параметрів без врахування розмітки даних порівняно з випадковою ініціалізацією, попереднім навчанням на наборі ImageNet та навчання в складні сегментатора зображення з архітектурою UNet. Отримана точність прийнятна до практичного використання, оскільки подібні помилки допускають і самі оператори. При цьому отримані результати перевершують результати аналогічних досліджень в рамках інших підходів [71, 98].

3.4 Оптимізація параметрів моделей розпізнавання дефектів труб водовідведення

Основною метою аналізу зображень є виявлення і локалізація дефектів. Для локалізації дефектів може бути використана інформація про контекст спостереження та значення одометру на кадрі зображення. Як було зазначено в підрозділі 2.1 аналіз дефектів труб в контексті “Камера спрямована вперед вздовж труби” слід виконувати не на всій фотографії, а на бокових частинах кадру, оскільки саме для них відомі значення з одометру, що спрощує їх локалізацію (рис. 3.13) [86, 19].

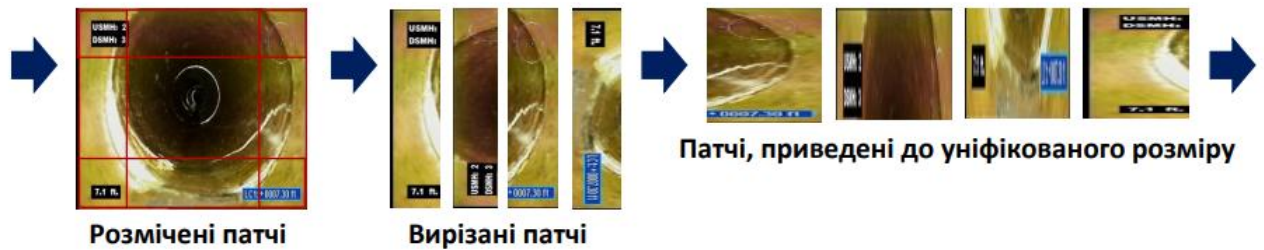


Рисунок 3.13 – Послідовність обробки розмічених патчів зображення дефектів, для яких відоме значення одометра

Аналіз рис. 3.13 показує, що під час формування навчального та тестового набору даних, зі звітних кадрів вирізаються прямокутні зображення, що відповідають 15% від зображення зверху, знизу та по бокам, а мітка формується за результатами парсингу звіту. Потім ці зображення приводяться до роздільної здатності 160x160 пікселів шляхом інтерполяції. Збір даних за умов навченого контексту можна автоматизувати, оскільки в архівних звітах може міститися необхідна інформація. Наприклад, у американських стандартах NASSCO (The National Social Safety-Nets Coordinating Office) та британських стандартах MSCC (Manual of Sewer Condition Classification) у звітах вказується момент відео, код дефекту та розміщення дефекту за годинниковою стрілкою, що дозволяє визначити на якому патчі знаходиться дефект [86]. Це простіше ніж виконувати розмітку обмежувальними прямокутниками для детектора [84]. Проте класифікатор контексту не є безпомилковим, а самі звіти мають неточності і суб'єктивну складову, що потребує залучення експерта для очищення набору набору даних і обмежує обсяги розмічених вибірок.

Алфавіт класів дефектів в контексті “Камера спрямована вперед вздовж труби” містить 9-ть найбільш поширених дефектів труб водовідведення. Клас X_1^o відповідає дефекту “roots”, що відповідає проникненню коренів дерев в трубу. Клас X_2^o відповідає дефекту “deposits”, що призводить до зменшення робочого перерізу труби водовідведення. Класи X_3^o , X_4^o , X_5^o , X_6^o та X_7^o характеризують різноманітні структурні дефекти і руйнування труби такі як “surface damage”, “crack”, “fracture”,

“broken” та “hole” відповідно. Клас X_8^o відповідає дефекту “infiltration”, який виникає внаслідок входження води через дефект або пошкоджений стик, або через пористу ділянку на стінці труби. Клас X_9^o відповідає дефекту “Obstruction”, що суттєво утруднює потік або зменшує гідравлічну потужність. X_{10}^o відповідає нормальному стану труби і необхідний для регуляризації.

Для кожного класу зібрано по 100 навчальних та тестових зразків. Дані для перевірки концепції надала компанія Molfar.AI sp. z o. o. (Гданськ, Польща), яка в свою чергу їх отримала від компанії Ace Pipe Cleaning, Inc (Канзас Сіті, США). Звичайно кожен з цих класів може мати численну кількість підкласів, які легко визначити на основі різноманітних логічних правил. Кожне зображення використовується для розширення навчального набору в 4 рази шляхом застосування невеликих випадкових викривлень, яка включає варіацію яскравості, масштабу та поворотів до 10% від повного діапазону, а також накладання шумів типу сіль та перець. Сформована вибірка розмічених даних є збалансованою і становить 7000 зображень, 15% з яких використано для тестування.

Для експериментів і перевірки концепції використано модель в рамках традиційного підходу (рис. 2.10а) та її модифікацію в рамках запропонованого підходу (рис. 2.10б). Як градієнтний метод оптимізації обрано адаптивний алгоритм Adam [19]. Зупинка навчання відбувається згідно механізму ранньої зупинки, тобто коли критерій на тестовій множині не покращується протягом 10 епох. Кожен міні-пакет даних містить 32 зображення, а коефіцієнт швидкості навчання (learning rate) дорівнює 0,0001.

Розглянемо результати машинного навчання класифікатора дефектів на патчах кадрів у контексті “Камера спрямована вперед вздовж труби” (forward) в рамках традиційного підходу. Оцінювання точнісних характеристик можна здійснювати з використанням таких критеріїв (criterion) як точність (2.4.6) та інформаційний критерій ефективності (2.4.16).

На рис. 3.14 показано зміни точності моделі на тестовій (test_acc) і навчальній (train_acc) вибірках залежно від номеру епохи навчання [19].

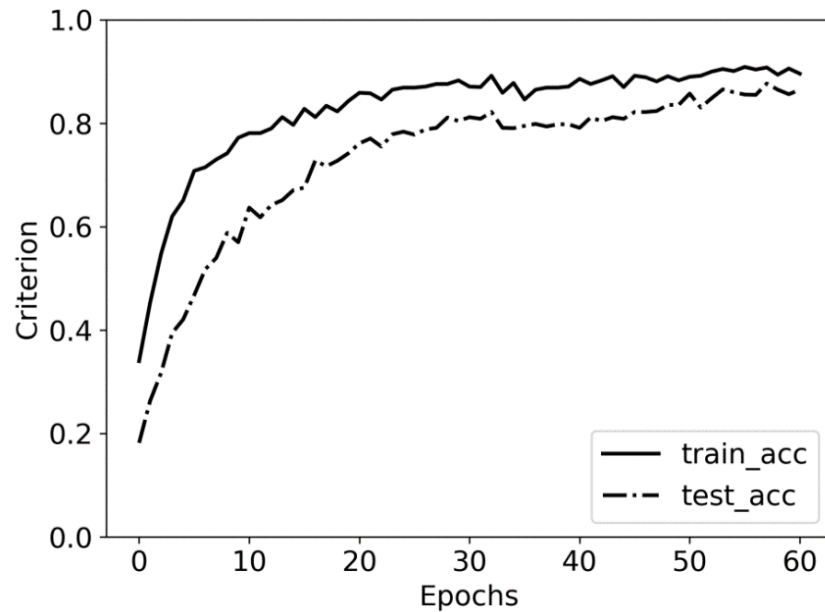


Рисунок 3.14 – Графік залежності точності (2.4.6) класифікації патчів кадру та функції втрат під час навчання класифікатора дефектів від кількості епох навчання в рамках традиційного підходу

Аналіз рис. 3.14 показує, що підвищення точності зупиняється після 40 епох навчання з максимальним значенням точності результуючої моделі, що дорівнює 86% на тестовій вибірці. При цьому відстань між кривою на тестовій і на навчальній вибірках свідчить про наявність невеликого ефекту перенавчання.

Розгляд результатів навчання в рамках запропонованого підходу почнемо з результатів другого етапу навчання з використанням запропонованої архітектури моделі (рис. 2.10б) і функції втрат (3.1.2) з регуляризацією (3.1.5). На рис. 3.15 показано залежність усереднених за алфавітом класів (2.4.20) значень точності (2.4.6) та інформаційного критерію (2.4.16) від кількості епох навчання (Epochs). При цьому побудова інформаційно-екстремальних класифікаційних вирішувальних правил (2.5.1) і обчислення інформаційного критерію (2.4.16) на тестовій (`test_info_cri`) та навчальній (`train_info_cri`) вибірках відбувається кожні 5 епох. Усереднене значення точності на тестовій (`test_acc`) та навчальній (`train_acc`) вибірках відбувається в кінці кожної епохи навчання. Після 30 епох навчання модель було збережено для наступного експерименту.

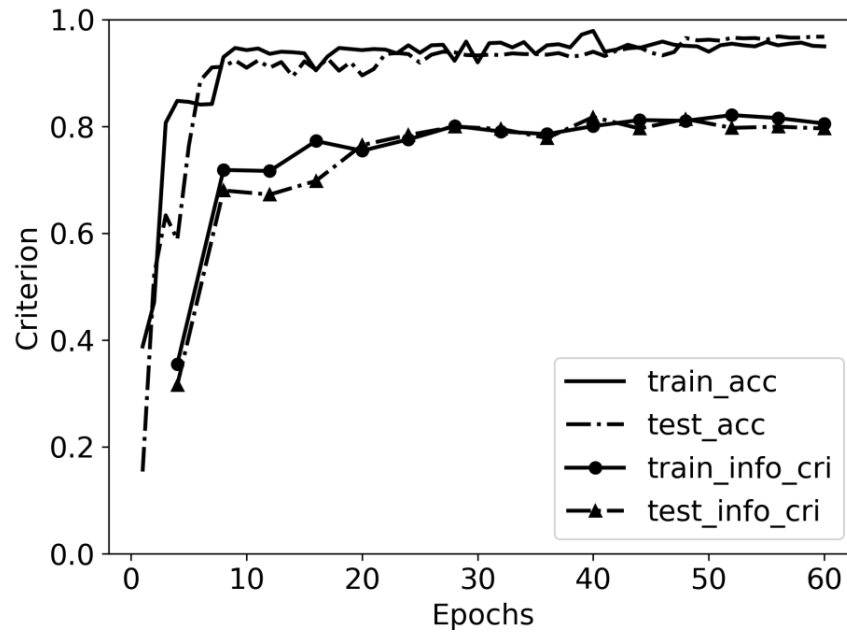


Рисунок 3.15 – Графік залежності усереднених значень точності (2.4.6) класифікації патчів кадру та інформаційного критерію (2.4.16) від кількості епох навчання в рамках запропонованого підходу з функцією втрат (3.1.2) та регуляризацією (3.1.5)

Аналіз рис. 3.15 показує, що з 10 епохи зростання точності значно уповільнилося, а після 40 епох було досягнуто максимальну точність, що дорівнює 95%, що на 9% більше ніж у випадку традиційного підходу. Однак після 40 епох навчання, подальше покращення практично припинилося. При цьому тестові і навчальні криві майже співпадають, що свідчить про високу узагальнюючу здатність отриманих вирішувальних класифікаційних правил (2.5.1). Усереднене за алфавітом класів значення інформаційного критерію дорівнює 0,815. Середня відстань між центрами гіперсферичних контейнерів дорівнює 19 кодових одиниць, а середній радіус контейнерів класів рівний 8 кодових одиниць. Це свідчить про високу завадозахищеність класифікаційних вирішувальних правил (2.5.1).

На рис. 3.16 показано подальші результати навчання моделі, збереженої на попередньому етапі після 30 епох навчання. Навчання здійснюється також 30 епох, але з використанням функції втрат (3.1.8) з метою зменшення внутрішньокласової дисперсії в просторі Хемінга [84].

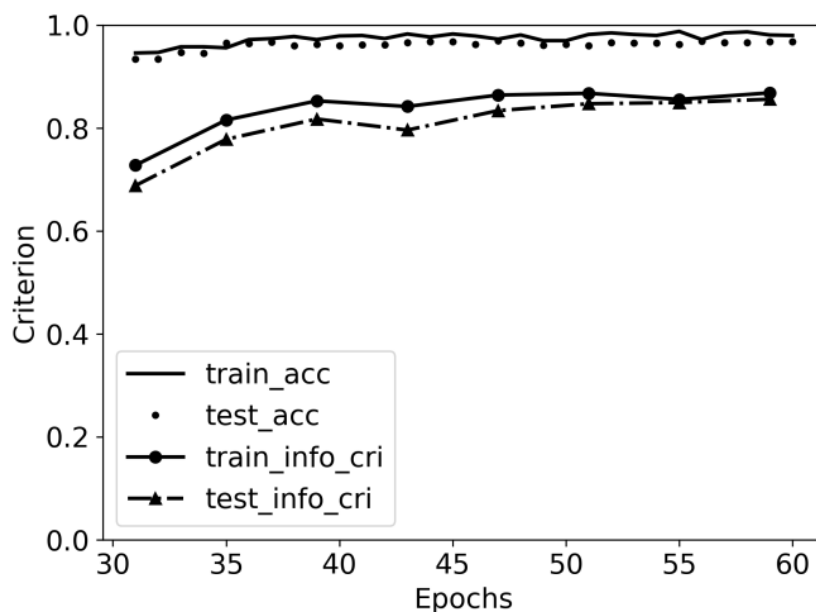


Рисунок 3.16 – Графік залежності усереднених значень точності (2.4.6) класифікації патчів кадру та інформаційного критерію (2.4.16) від кількості епох навчання в рамках запропонованого підходу з функцією втрат (3.1.8)

Аналіз рис. 3.16 показує, що побудова інформаційно-екстремальних класифікаційних вирішувальних правил (2.5.1) і обчислення інформаційного критерію (2.4.16) на тестовій (test_info_cri) та навчальній (train_info_cri) вибірках відбувається кожні 5 епох. При цьому усереднене значення точності на тестовій (test_acc) та навчальній (train_acc) вибірках відбувається в кінці кожної епохи навчання. Аналіз результатів показує, що використання взаємної бінарної кросентропійної функції втрат (3.1.8) забезпечує зростання точності на 3% порівняно з попереднім етапом і становить 98%. При цьому усереднене за алфавітом класів значення інформаційного критерію (2.4.16) дорівнює 0,856.

Таким чином мультифазне навчання з інформаційно-екстремальними вирішувальними правилами забезпечує підвищення точності розпізнавання дефектів у контексті “орієнтація камери вперед вздовж труби” порівняно з традиційним підходом на 11% за туж кількість епох навчання за умов обмеженого обсягу розмічених навчальних даних.

В більшості стандартів протокол інспекції передбачає, що оператор кравлера повертає камеру для детальнішого огляду дефекту і оцінювання його тяжкості,

тому розглянемо задачу розпізнавання дефектів при повороті камери на стінку труби. Для підготовки вхідного математичного опису класифікатора кадрів у контексті повороту камери до стінки труби було використано SEWER-ML набір даних [101]. Набір даних містить 1,3 мільйони зображень з 75618 відеофайлів, зібраних з датських компаній, що обслуговують трубопроводи водовідведення. Проте датський стандарт, за яким здійснена розмітка вибірки, не дозволяє використати розмітку для конвертації в інші стандарти, в той час як інші стандарти можуть бути конвертовані в датський стандарт. Тому здійснена ручна перерозмітка даних. Було виділено 13 класів розпізнавання, 12 з яких показано на рис. 3.17 [79].

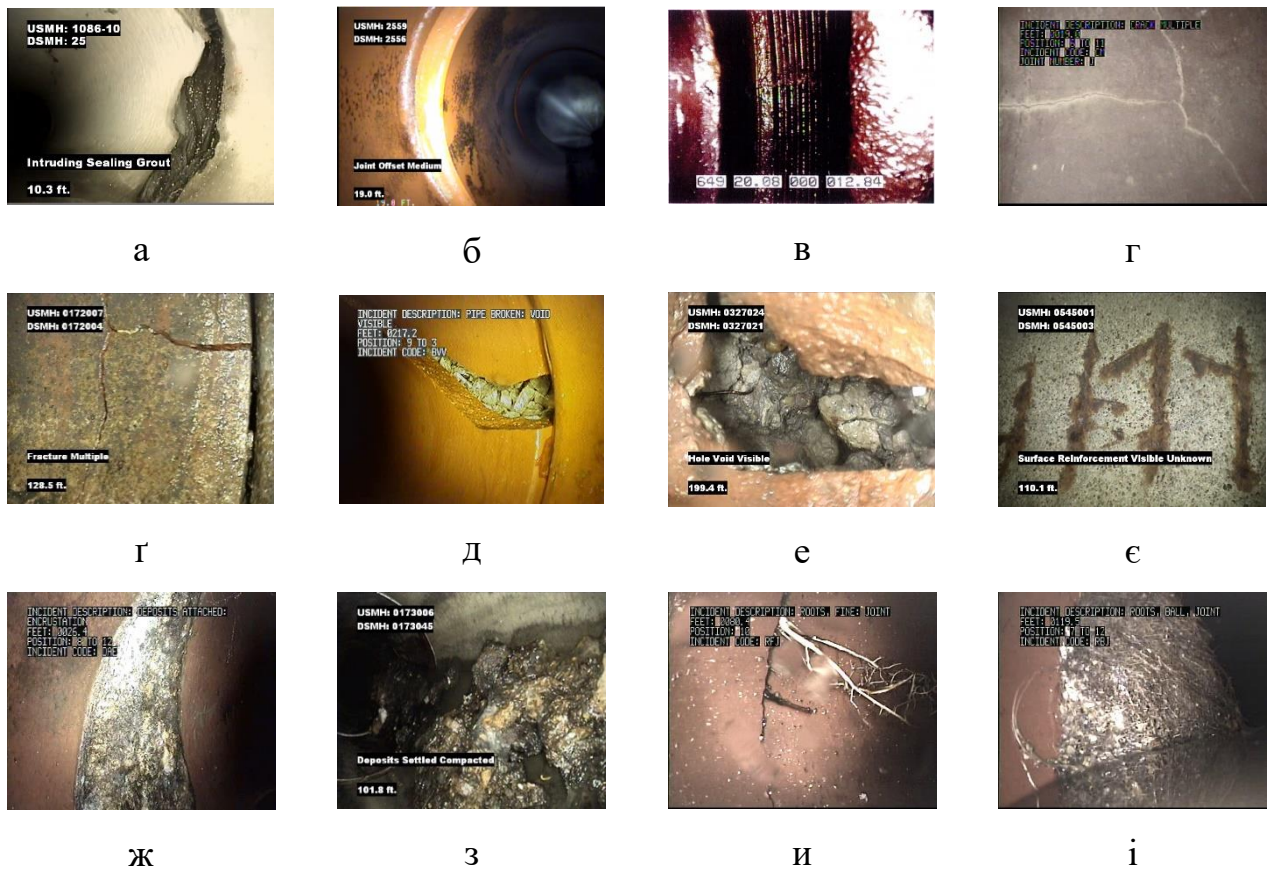


Рисунок 3.17 – Приклади зображень кожного з класів дефектів у контексті “Камера спрямована на стінку труби ”: а – клас X_1^0 ; б – клас X_2^0 ; в – клас X_3^0 ; г – клас X_4^0 ; д – клас X_5^0 ; е – клас X_6^0 ; ж – клас X_7^0 ; з – клас X_8^0 ; и – клас X_9^0 ; і – клас X_{10}^0 ; і – клас X_{11}^0 ; і – клас X_{12}^0

Для задачі попереднього навчання без вчителя планується використати 20% набору даних. А для точної настройки під задачу розпізнавання дефектів планується обрати тільки ті зображення, що містять дефекти в контексті “вид збоку”. При цьому зображення приводяться до роздільної здатності 224x224 пікселів з метою зменшення витрат ресурсів на навчання та екзамен (inference).

Опис алфавіту сформованих класів показано в табл. 3.4. Поділ вибірки на тестову і навчальну здійснюється у відношенні 70% на навчання і 30% на тестування. При цьому під час поділу використовується стратегія збереження пропорції кількості зразків кожного класу в тестовій і навчальній вибірці.

Таблиця 3.4 – Структура розміченої вибірки даних для розпізнавання дефектів на стінках труб водовідведення

Позначення класу	Назва дефекту	Кількість розмічених зразків
X_1^o	Порушення цілісності прокладки на стику	250
X_2^o	Відкритий стик	200
X_3^o	Зміщений стик	331
X_4^o	Тріщина	270
X_5^o	Розлом	189
X_6^o	Розбита труба	306
X_7^o	Дірка в трубі	450
X_8^o	Пошкодження поверхні	210
X_9^o	Прикріплені відкладення	200
X_{10}^o	Осад	300
X_{11}^o	Тоненькі корені	500
X_{12}^o	Значне розростання коренів	310
X_{13}^o	Нормальний стан труби	500

Балансування навчальної підвибірки здійснюється за рахунок аугментації класів з меншою кількістю зразків. Аугментація полягає у застосуванні операцій масштабування (збільшення з обрізанням) до 5%, обертання з обрізанням на ± 18 градусів та збільшенням чи зменшенням яскравості до 5%. Розмір навчальної вибірки кожного класу після аугментації для класифікатора дефектів становить 500 зразків.

На рис. 3.18 показано графік зміни функції втрат (3.1.1) від кількості ітерацій (Number of iterations) [79]. Процес оптимізації зупиняється, якщо функція втрат суттєво не змінюється протягом 10 епох. Однак конкретне значення функції втрат не є інформативним, більш цікавим є кількість ітерацій навчання до зупинки, щоб оцінити час і ресурси необхідні для ініціалізації ваг в процесі самонавчання. Аналіз рис. 3.18 показує, що після 6 000 ітерацій функція втрат (3.1.1) помітно не змінювалась. При цьому розмір міні-паketу становить 512 зображень.

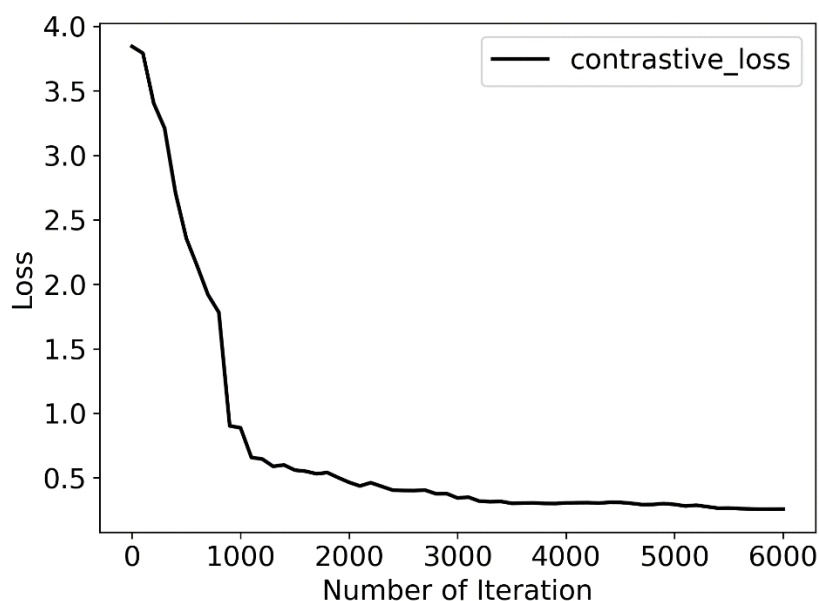


Рисунок 3.18 – Графік залежності контрастної функції втрат (3.1.1) від кількості ітерацій навчання на міні-пакетах нерозміченого набору даних Sewer-ML

Вибірка розмічених даних незбалансована, тому для оцінювання ефективності моделі класифікатора пропонується використати мікроусереднене

значення (2.4.19) критерію F1 (2.4.7). На рис. 3.19 показано графік зміни мікроусередненого за алфавітом класів значення метрики F1 (2.4.7) на тестовій вибірці (F1_val) під час навчання класифікатора дефектів за модифікованою триплетною функцією втрат NCA (3.1.2) (F1_val_nca) та регуляризацією (3.1.5) [79].

Показана на рис. 3.19 крива F1_val_nca_pretrained показує залежність усередненої метрики F1 (2.4.7) на тестовій вибірці від кількості епох навчання з функцією втрат (3.1.2) та регуляризацією (3.1.5), але після контрастного самонавчання з екземпляр-прототипною контрастною функцією втрат (3.1.1).

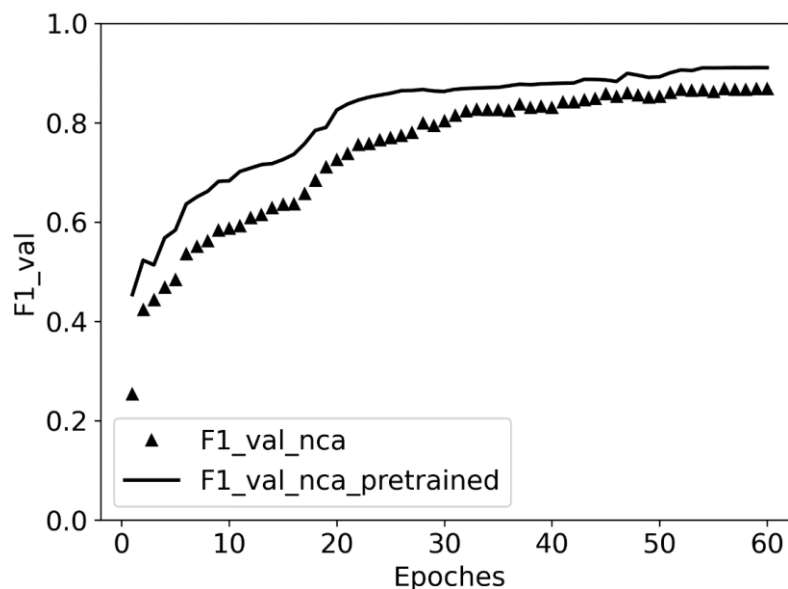


Рисунок 3.19 – Графік залежності значення мікроусередненого за алфавітом класів значення метрики F1 (2.4.7) на тестовій вибірці від кількості епох навчання за функцією втрат (3.1.2) з регуляризацією (3.1.5) до і після попереднього навчання без вчителя з використанням функції втрат (3.1.1)

Аналіз рис. 3.19 показує, що попереднє навчання без вчителя покращує точнісні характеристики моделей, що навчаються як за функцією втрат (3.1.2) з регуляризацією (3.1.5) [79]. При цьому результат навчання без попереднього самонавчання забезпечує значення метрики $F1=0,911$, а результат навчання після попереднього самонавчання забезпечує значення метрики $F1=0,927$.

Для наступного етапу запропонованого методу використаємо моделі з переднього етапу навчання, заморожені після 30 епох, як стартові точки пошуку. Для порівняння здійснимо навчання моделей з традиційною структурою вихідного шару (рис. 2.10a) і традиційним методом оптимізації. На рис. 3.20 показано графік зміни мікроусередненого значення за алфавітом класів метрики F1 (2.4.7) на тестовій вибірці ($F1_val$) під час навчання класифікатора дефектів за традиційним методом без попереднього самонавчання ($F1_val_baseline$) та з попереднім контрастним самонавчанням ($F1_val_baseline_pre$) [79]. Крім цього на графіку показано зміни мікроусередненого значення за алфавітом класів метрики F1 (2.4.7) на тестовій вибірці під час IV етапу навчання (рис. 3.1) з крос-ентропійною функцією втрат (3.1.8) для моделей, що були попередньо навчені під час самонавчання та з учителем ($F1_val_nca_pretrained$) за модифікованою триплетною функцією втрат NCA (3.1.2) з регуляризацією (3.1.5).

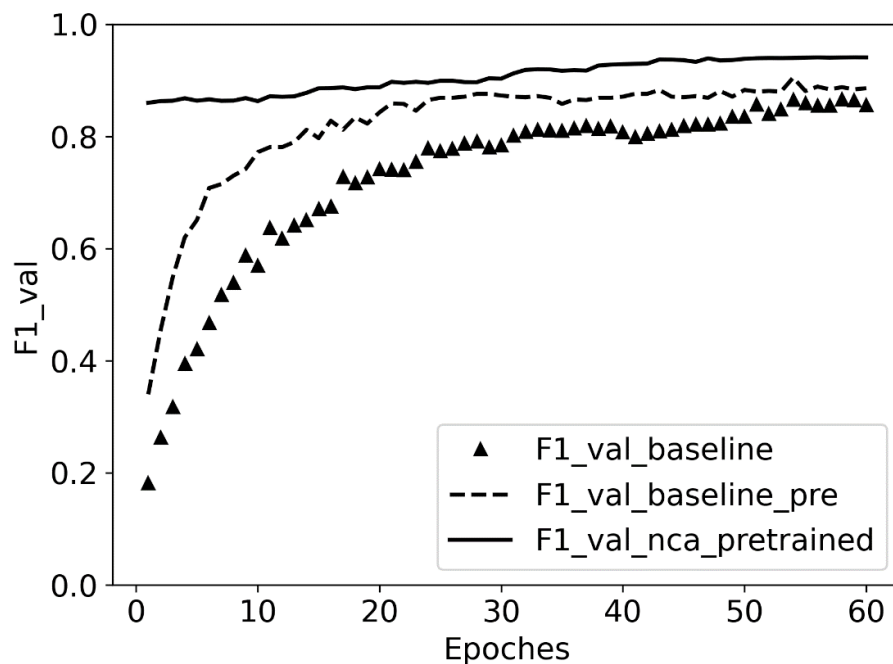


Рисунок 3.20 – Графік залежності значення мікроусередної за алфавітом класів значення метрики F1 (2.4.7) на тестовій вибірці від кількості епох навчання за традиційним методом за запропонованим методом навчання

Аналіз рис. 3.20 показує, що запропонований метод навчання має переваги у порівнянні з традиційним методом навчання за метрикою F1(2.4.7). Запропонований метод навчання дозволяє підвищити метрику F1 на 5% порівняно з традиційним підходом. В результаті оптимальна модель класифікатора дефектів на стінках труб водовідведення має таке значення метрики $F1=0,977$, що свідчить про високу достовірність вирішувальних правил.

У праці [101] розглядалась задача класифікації дефектів на датасеті SewerML, де результати оцінювалися за F1-метрикою. Отримане значення F1-міра не перевищує 0,9108. Коди алфавіту класів SewerML згруповані по іншому, що не дозволяє на пряму співставляти з нашим алфавітом, однак дає змогу оцінити точність результуючих звітів, хоч і в іншому стандарті. Звідси можна зробити висновок, що результати авторів поступатимуться результатам запропонованого підходу.

Таким чином, експериментально доведено перевагу використання запропонованої моделі та багатофазного методу навчання для класифікаційного аналізу дефектів на стінках труб водовідведення. Показано, що використання попереднього контрастного самонавчання покращує результати як традиційного, так запропонованого методу навчання. Отримана точність прийнятна до практичного використання, оскільки подібні помилки допускають і самі оператори.

Висновки до третього розділу

У третьому розділі розроблено алгоритми оптимізації параметрів функціонування здатної навчатися системи відеомоніторингу труб водовідведення та досліджено вплив архітектурних рішень і гіперпараметрів на функціональну ефективність машинного навчання. Одержано такі наукові результати.

1. Розроблено метод навчання моделі класифікаційного аналізу даних, що полягає в реалізації наступних етапів :

- ініціалізація вагових коефіцієнтів за принципами переносу знань та навчання без вчителя з використанням екземпляр-прототипної контрастної функції втрат;
- уточнення вагових коефіцієнтів на розміченій вибірці з використанням NCA функції (Neighbourhood Components Analysis) втрат з майнінгом триплетів і регуляризації для штрафування за помилку дискретизації ознакового подання;
- визначенням кодового подання кожного з класів;
- фінальне уточнення вагових коефіцієнтів з використанням бінарної ентропійної функції втрат та отриманого кодового подання цільових міток класів;
- пошук оптимальних за інформаційним критерієм кодових радіусів радіально-базисних класифікаційних вирішувальних правил в двійковому просторі Хемінга;
- зміна гіперпараметрів і повторення всіх кроків для пошуку оптимальної в інформаційному сенсі конфігурації моделі.

2. Розроблено алгоритм навчання класифікатора контексту спостереження з урахуванням сусідніх кадрів для задачі аналізу даних відеоінспекції труб водовідведення. Для цього спочатку навчається класифікатор контексту окремих кадрів, а потім отримана модель використовується для формування ознакового опису поточного і сусідніх кадрів, що аналізуються темпоральною нейронною мережею. Встановлено перевагу використання згорткової моделі порівняно з LSTM та GRU мережами для аналізу контексту спостереження та запропонованого методу багатоетапного навчання. Також емпірично доведено, що регулярна згорткова архітектура темпоруальної мережі є найбільш ефективною під час аналізу міжкадрових залежностей відеоінспекції труб. Для алфавіту з 11 класів контексту навчена за традиційним алгоритмом модель з регулярною темпоруальною згортковою мережею забезпечує такі мікроусереднені за алфавітом класів значення F-міри, точності та інформаційного критерію як 0,87, 88,1% та 0,57 відповідно. При цьому навчання в рамках запропонованого підходу забезпечує більші мікроусереднені значення F-міри, точності та інформаційного критерію, що становлять 0,92, 93% та 0,67 відповідно.

3. Запропоновано модель і алгоритм навчання класифікатора рівня води на зображеннях відеоінспекції труб водовідведення з урахуванням сусідніх кадрів. Доведено перевагу використання запропонованої моделі для аналізу рівня води і запропонованого багатоетапного методу її навчання. Експериментально показано, що найбільш ефективною моделлю для класифікаційного аналізу рівня води на відеокадрах є комбінація згорткової 2D-мережі і темпоральної згорткової мережі з регулярними згортками. При цьому згорткова темпоральна мережа продемонструвала вищу ефективність врахування міжкадрових залежностей порівняно з рекурентними мережами. Також продемонстровано перевагу у використанні екземпляр-прототипної контрастної функції втрат для ініціалізації параметрів без врахування розмітки даних порівняно з випадковою ініціалізацією, попереднім навчанням на наборі ImageNet та навчання в складні сегментатора зображення з архітектурою UNet. Для алфавіту класів, що описує рівнів води від 0% до 55% з кроком в 5% мікроусереднене значення F-міри становить 0,879. Крім того отримані результати перевершують результати аналогічних досліджень інших авторів в рамках інших підходів.

4. Розроблено модель і алгоритм навчання класифікатора дефектів з урахуванням контексту спостереження. Для контексту спостереження, що відповідає орієнтації камери вперед вздовж труби здійснюється аналіз класифікаційний аналіз бічних патчів зображення, що дозволяє визначити повздовжню та радіальну локалізацію дефекта. Для контексту спостереження, що відповідає орієнтації камери на бічну стінку здійснюється класифікаційний аналіз цілого зображення, повздовжня інформація уточнюється по одометру, а радіальна локалізація по попереднім значенням контексту спостереження. Експериментально доведено, що використання попереднього контрастного самонавчання покращує результати як традиційного, так запропонованого методу навчання. При цьому запропонований багатоетапний метод машинного навчання забезпечує вищу на 11% точність класифікаційного аналізу дефектів у контексті спостереження, що відповідає орієнтації камери вперед вздовж труби, порівняно з традиційним підходом і становить 98%. Оскільки в контексті спостереження, що відповідає

орієнтації камери на бічну стінку, присутня помітна незбалансованість даних, то ефективність класифікації дефектів оцінюється F1-метрикою і становить $F1=0,977$, що на 5% більше ніж в рамках традиційного підходу. Отримана точність класифікації дефектів в рамках запропонованого підходу перевищує на 6% результати отримані іншими авторами, і дозволяє отримати більшу точність звітів про інспекцію.

Основні наукові результати, наведені у третьому розділі, опубліковано у працях автора [19], [22], [23], [79], [80], [81], [84], [85], [86], [89], [93], [94], [100],

РОЗДІЛ 4

ІНФОРМАЦІЙНЕ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ ОЦІНЮВАННЯ ФУНКЦІОНАЛЬНОГО СТАНУ ТРУБОПРОВОДУ ВОДОВІДВЕДЕННЯ

4.1 Структура та функції інтелектуальної системи оцінювання функціонального стану труб водовідведення за даними відеоінспекції

На основі отриманих в дисертаційній роботі наукових результатів розроблено структурні та функціональні схеми, інформаційне та програмне забезпечення інтелектуальної системи аналізу даних відеоінспекції труб водовідведення. Дана система здійснює розпізнавання рівня води, дефектів та їх локалізацію за рахунок класифікаційного аналізу відеозображення чи окремих його ділянок. Інформаційна технологія, що реалізує дану систему, включає як процес синтезу моделей аналізу даних, так і процес обробки інформації під час розпізнавання, тобто безпосереднього прийняття рішень.

Пропонована інформаційна технологія синтезу моделей аналізу даних передбачає наявність розмічених та нерозмічених навчальних даних, обрану архітектуру екстрактора ознак, розроблений алгоритм машинного навчання та бібліотечні класи і функції. Узагальнена функціональна модель процесу оптимізації параметрів системи аналізу даних в рамках запропонованої інформаційної технології показано на рис. 4.1 [102]. Використання кодових бібліотек глибоких нейронних мереж, обробки багатовимірних масивів та машинного зору дозволяють скоротити час розроблення і розгортання системи. Під обраною архітектурою екстрактор ознак маємо на увазі і основні його гіперпараметри, що впливають на інформаційну ємність та узагальнюючу здатність моделі. Синтез моделі класифікаційного аналізу здійснюється в процесі машинного навчання, деталізована схема якого показано на рис. 4.2.

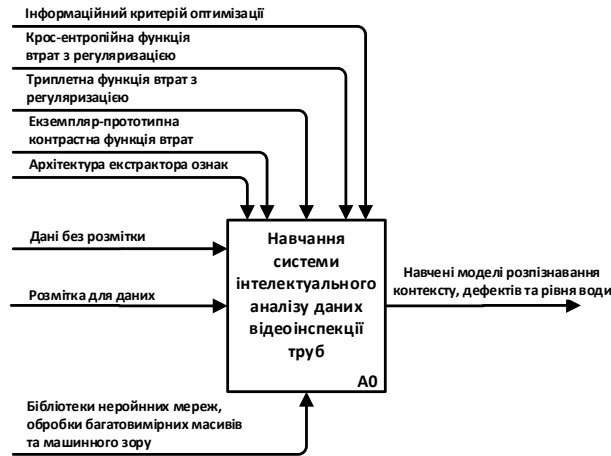


Рисунок 4.1 – Узагальнена функціональна модель процесу оптимізації параметрів системи аналізу даних в рамках запропонованої інформаційної технології

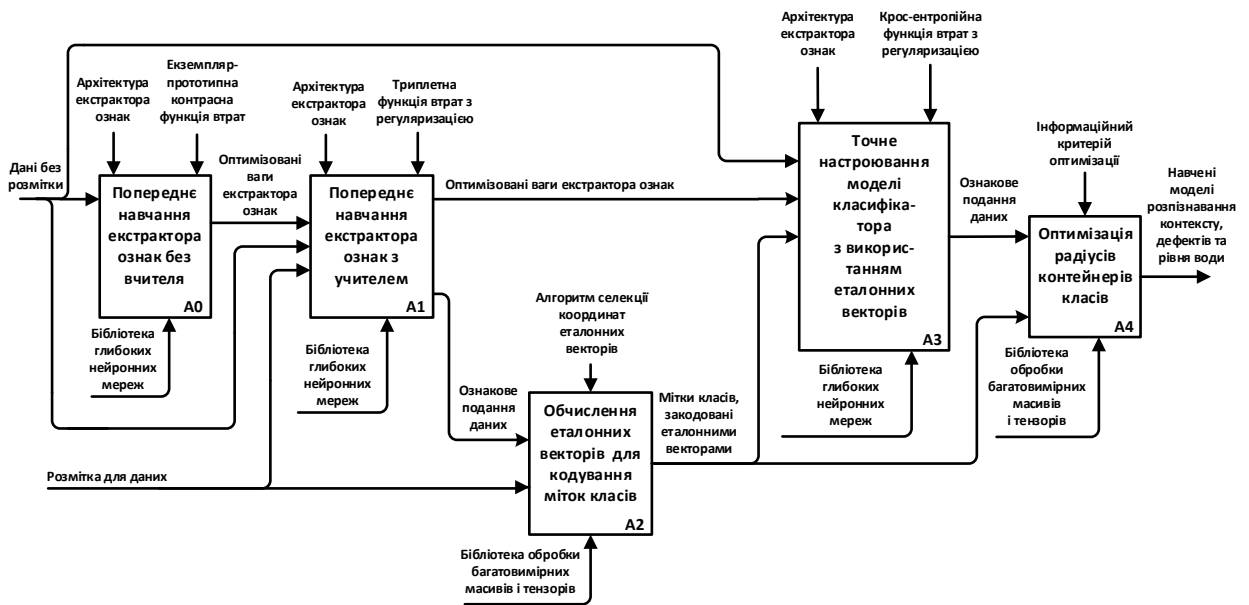


Рисунок 4.2 – Деталізована функціональна модель процесу оптимізації параметрів системи аналізу даних в рамках запропонованої інформаційної технології

Аналіз рис. 4.2 показує, що процес синтезу моделі супроводжується зміною функції втрат, переходом від нерозмічених даних до розмічених, заміною міток класів на відповідні коди, а на останньому етапі і заміною оригінальних даних їх двійковим ознаковим поданням. Для формування кодів класів необхідно обрати

один з алгоритмів селекції координат еталонних векторів. Це можна зробити або усередненням ознакового подання кожного класу, або за формулою (3.1.7). При цьому структурна схема процесу синтезу моделі для задачі розпізнавання рівня води, контексту і дефектів в різних контекстах однакова, змінюватимуться лише навчальні дані і гіперпараметри моделі.

У робочому режимі системи відеомоніторингу необхідно об'єднати роботу всіх моделей і сформуванати єдиний результат у вигляді звіту про інспекцію (рис. 4.3). На рис. 4.4 показано деталізовану функціональну модель процесу аналізу даних в рамках інформаційної технології [102].

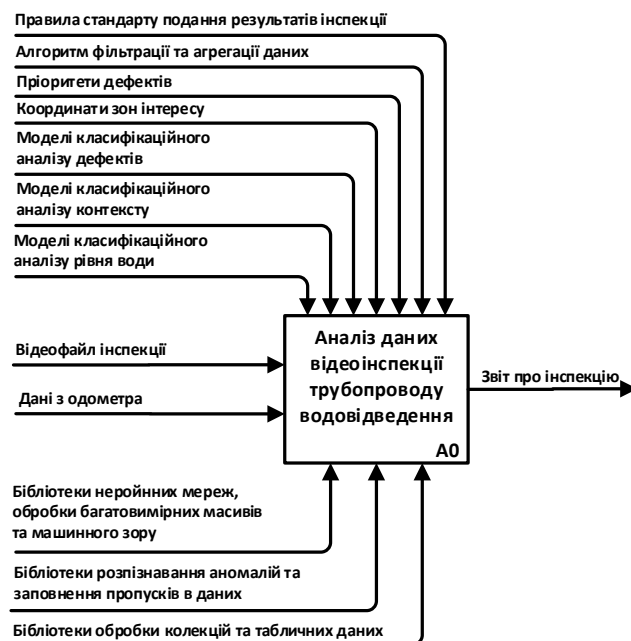


Рисунок 4.3 – Узагальнена функціональна модель процесу аналізу даних в рамках запропонованої інформаційної технології

Аналіз рис. 4.4 показує, що для аналізу рівня води та контексту спостереження використовуються як моделі, що не аналізують міжкадрові залежності, так і моделі, що уточнюють результат розпізнавання на кадрі шляхом одночасного аналізу сусідніх кадрів у певному часовому вікні. При цьому дефекти агрегуються окремо для різних контекстів спостереження. Інформація про рівень води потрібна для відмітки моментів зміни рівня води більше ніж на 10%, що

вимагає американський та британський стандарт [103]. Крім того, якщо рівень води збільшився більше ніж на 10%, а потім відновився, то це кодується як дефект просадки труби. Однак, якщо зміна рівня води була викликана перешкодами на шляху течії води, то це не вважається просадкою труби.

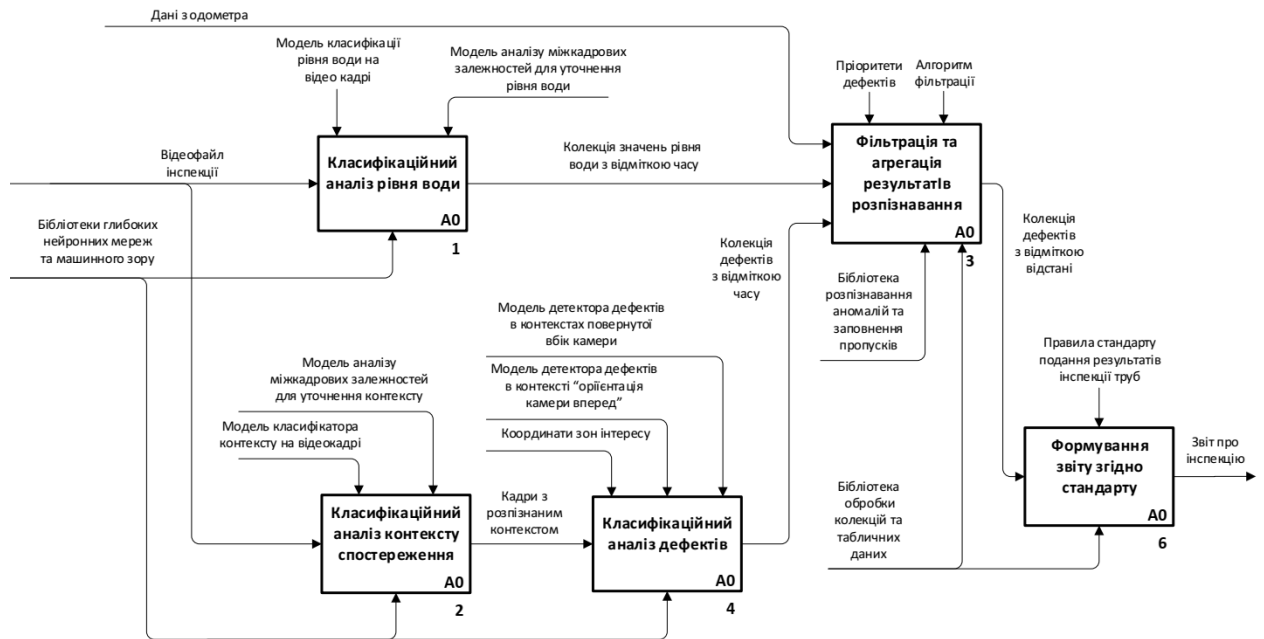


Рисунок 4.4 – Деталізована модель процесу оцінювання функціонального стану труб в рамках запропонованої інформаційної технології

Під час фільтрації дефектів результати розпізнавання агрегуються у пакети кадрів, у кожному з яких незмінне значення одометра. В рамках кожного пакету структурні дефекти з більшою тяжкістю (пріоритетом) в однакових ділянках труби можуть поглинати дефекти з меншою тяжкістю (пріоритетом). Структурні дефекти у порядку зростання тяжкості (пріоритету) утворюють певну послідовність розвитку uszkodження труби: пошкодження поверхні; тріщина; розлом; розбиття; дірка; деформація; руйнування або завал труби. Якщо у пакеті кадрів з однаковим значенням одометра список дефектів в одній і тій же частині труби розпізнаються одночасно, але не є сумісними, то обирається той код дефекту, що з'являється частіше.

Для ефективної агрегації дефектів необхідно локалізувати дефект як по дистанції так і по радіальних координатах (частина труби : верх, низ, ліво, право). Дані з одометра можуть перериватися чи видавати хибні значення внаслідок апаратних чи програмних збоїв, тому необхідно використовувати алгоритми виявлення аномальних значень та заповнення пропусків. Аномальні значення видаляються і стають пропусками. Пропуски заповнюються на основі методу найближчих сусідів, а пошук аномалій здійснюють на основі обчислення локального фактора викиду (LOF) кожного об'єкта, що передбачає порівняння локальної щільності розподілу в даній точці з щільністю розподілу в сусідній точці даних [104].

Для обслуговування максимальної кількості клієнтів, що можуть бути розподілені по всьому світу, система проектувалася для розгортання в хмарному середовищі. Для цього було розроблено веб-сервер, що приймає через веб-інтерфейс відеодані з записом значень одометра та описом параметрів інспекції. Відеодані зберігаються у хмарному сховищі файлів, а решта інформації записується до бази даних. Обробка кожного відео відбувається паралельно в сервісах пакетної обробки даних або з використанням безсерверних лямда-обчислень. Результати зберігаються також в базу даних, а згенеровані файли звітів у сховище файлів. Асинхронність процесів надсилання задач у обробку та отримання статусу і проміжних результатів реалізується за рахунок черг повідомлень. Узагальнену архітектуру системи показано на рис. 4.5.

На рис. 4.5 показано, що основними задачами пакетної обробки є розпізнавання контексту, дефектів та рівня води, а також фільтрації та агрегації результатів розпізнавання для формування звітів. Для формування звіту по конкретній інспекції потрібно здійснити послідовно обробку всіх задач пакетної обробки відповідного відео. Послідовність обробки регулюється за рахунок поля статусу в чергах повідомлень. Процес аналізу даних та формування звіту, що відбувається під час пакетної обробки, вже було описано схемою на рис. 4.4.

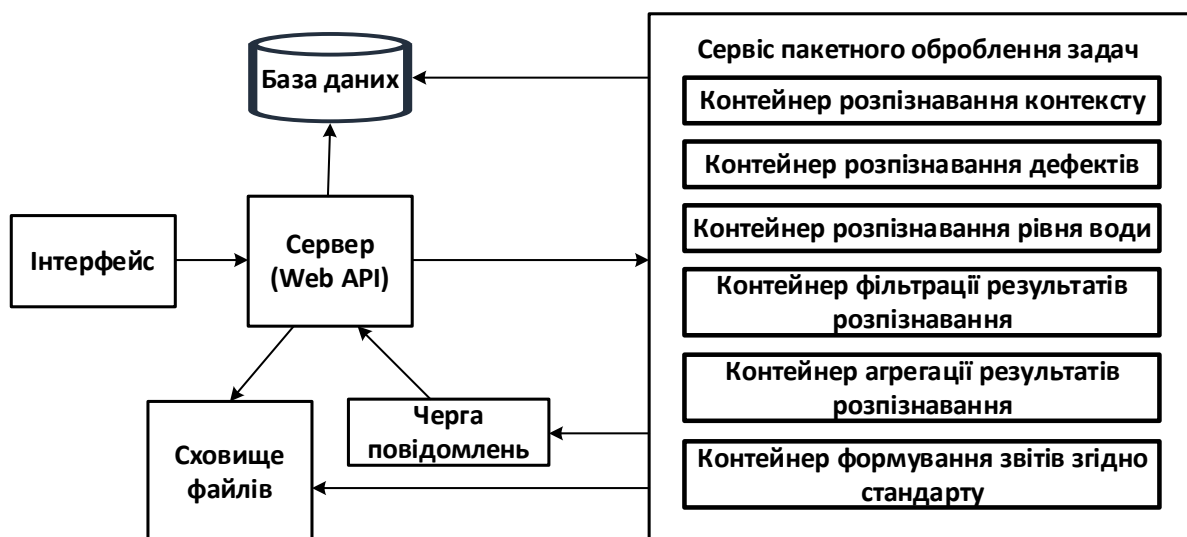


Рисунок 4.5 – Узагальнена структурна схема хмарного сервісу аналізу даних відеоінспекції труб водовідведення

Таким чином, аналіз даних відеоінспекції труб водовідведення має багатоетапну структуру як в режимі навчання, так і в режимі екзамону. При цьому в режимі синтезу кожна модель проходить етапи запропонованого методу навчання, а режимі екзамону відео в певній послідовності аналізується різними моделями, а результати аналізу фільтруються, агрегуються і кодуються згідно стандарту. Такий спосіб обробки зручно реалізується в хмарному середовищі за рахунок черг повідомлень і пакетної обробки задач.

4.2 Формування звітів про функціональний стан трубопроводу водовідведення

Одним з перших етапів обробки відео є фільтрація і відновлення даних від одометра, які вказують на пройденою камерою відстань в трубі від початкового до кінцевого колодязя. Внаслідок апаратно-програмних збоїв виникають викиди і пропуски в даних. В деяких апаратно-програмних реалізаціях дані одометра доступні тільки у вигляді анотацій на відео-кадрі, тому їх зчитування потребує використання алгоритмів розпізнавання тексту. Найкращі представники цих алгоритмів основані на поєднанні OCR-дескрипторів та глибоких нейронних мереж

[105]. Незважаючи на те, що подібні алгоритми досягли якості розпізнавання на рівні людини, внаслідок впливу шуму та артефактів відеозображення можуть виникати помилки розпізнавання, що так само розглядається як викиди та пропуски в даних. На рис. 4.6 показано типовий приклад зміни значень одометра в часі (залежно номера кадру) під час відеоінспекції труб.

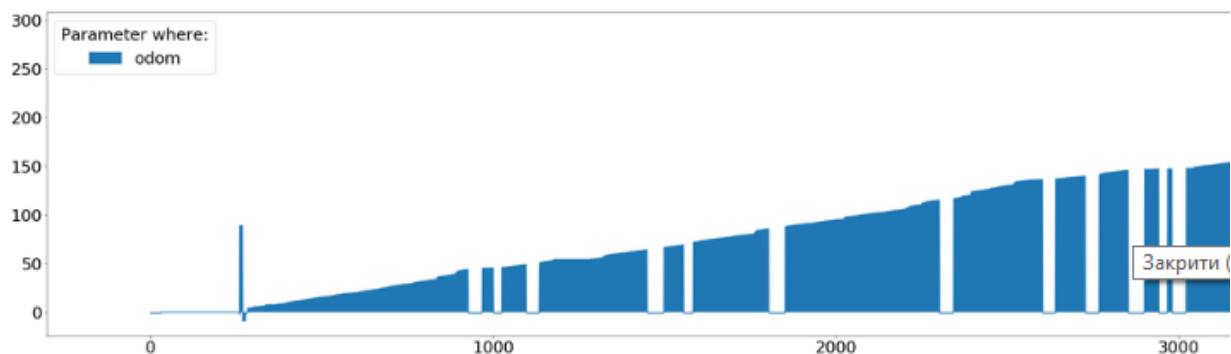


Рисунок 4.6 – Приклад серії значень одометра, які змінюються в часі протягом інспекції труб водовідведення

Для виявлення і виправлення в даних одометра викидів та пропусків можна використовувати локальний фактор викиду (LOF) та метод найближчого сусіда. Однак більш ефективним методом виявлення викидів вважається використання фільтра Хампеля [106]. Пропуски і викиди в даних можна замінити нульовим значенням для подальшого відновлення. При цьому для виявлення тривалих пропусків в даних можна ефективно використовувати апріорну інформацію про те, що камера рухається лише в одному напрямку – вперед від початкового колодязя до кінцевого колодязя. Викиди і пропуски можна відновити шляхом інтерполяції. На рис. 4.7 показано результат фільтрації і відновлення даних одометра, проілюстрованих на рис. 4.6.

Аналіз рис. 4.7 показує, що вдалося позбутися викиду значень одометра, що стався приблизно на 300 секунді інспекції і відновити пропуски, які з'являються пізніше. Отриману послідовність даних можна використати для згортки результатів розпізнавання контексту і дефектів на кадрах, що відповідають однакової відстані.

Під час згортки результатів розпізнавання дефектів використовуються результати розпізнавання контексту, який надає інформацію про локалізації дефектів в радіальних координатах. На рис. 4.8 показана поширена у Польщі форма опису локалізації дефектів на 4-ри основні позиції або їх комбінації, коли спостереження на межі двох класів [107].

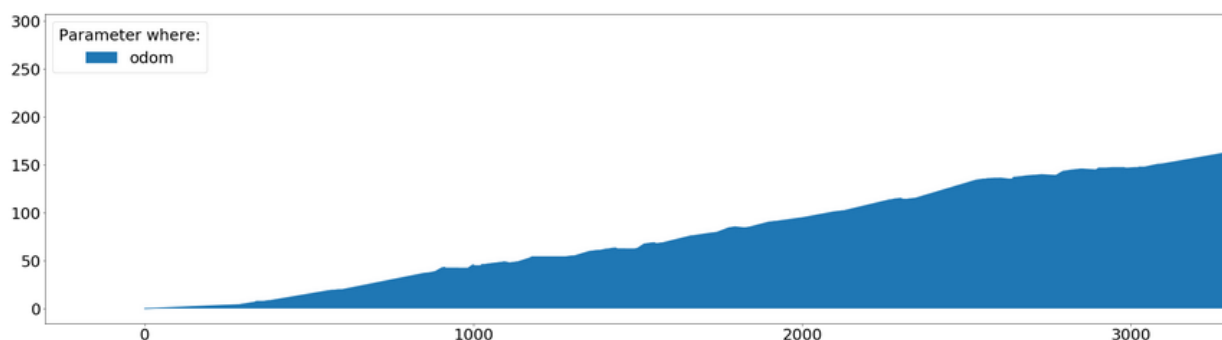


Рисунок 4.7 – Приклад відфільтрованих і відновлених покадрових даних одометра, отриманих під час інспекції труб водовідведення

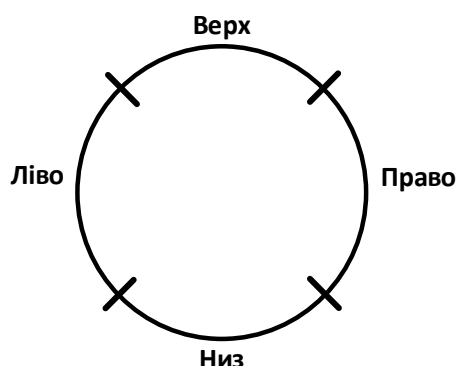


Рисунок 4.8 – Контур труби і найменування сегментів, що вказується у звітах для локалізації розпізнаних дефектів

У випадку складних зашумлених зображень низької роздільної здатності відбувається зниження точності розпізнавання контексту спостереження, однак це компенсується в процесі фільтрації і згортки [78]. Приклад результатів покадрового розпізнавання контексту на зашумлених даних відеоінспекції без фільтрації показано на рис. 4.9.



Рисунок 4.9 – Приклад результату покадрового розпізнавання контексту спостереження на даних відеоінспекції труб водовідведення

Під час фільтрації результатів розпізнавання контексту враховується, що колодязі знаходяться по краях труби, а посередині їх не повинно бути, тому відповідні спрацювання будуть заміщуватися спрацюваннями на інші класи, що знаходяться по сусідству. Крім того враховується апріорна інформація про те, що контекст не змінюється миттєво, тому у вікні 1-2 секунди можна здійснювати згортку, тобто обирати клас розпізнавання, що найчастіше зустрічається. Так само можна порівнювати кадри між собою, якщо вони мало відрізняються, то результати розпізнавання на цих кадрах також можуть підлягати згортці. Для порівняння можна використовувати середньоквадратичне відхилення, яке повинно не перевищувати 0,001. Результат фільтрації даних покадрового розпізнавання контексту, показаних на рис. 4.9, показано на рис. 4.10.

Аналіз рис. 4.10 показує, що камера під час інспекції більшу частину шляху орієнтована вперед вздовж труби і лише інколи повертається для уточнення інформації.

Аналогічно до обробки результатів розпізнавання контексту спостереження здійснюється обробка результатів розпізнавання рівня води [108, 100]. При цьому врахування апріорної інформації про неможливість миттєвої зміни рівня води можуть бути використана фільтрація для кадрів з однаковим значенням відстані та для кадрів у вікні 1-2 метри. Функція фільтрації реалізована у вигляді заміни класу

розпізнавання на конкретному кадрі класом, що найчастіше зустрічається серед сусідніх кадрів з однаковим значенням відстані чи в рамках вікна 1-2 метри.

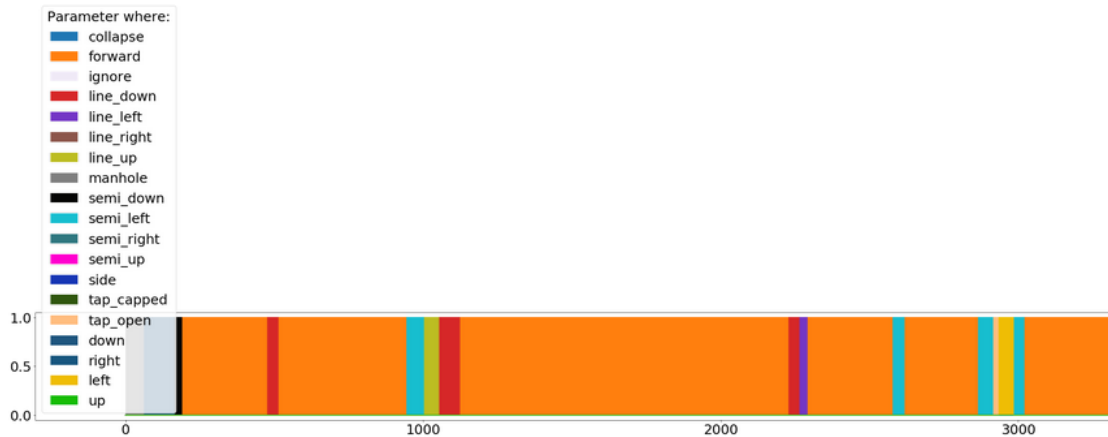


Рисунок 4.10 – Приклад результату покадрового розпізнавання контексту спостереження на даних відеоінспекції після фільтрації

При цьому під час агрегації результатів можна враховувати контекст розпізнавання шляхом ігнорування результатів розпізнавання на кадрах, що не відповідають контексту орієнтації камери вперед вздовж труби. Кадри з іншим контекстом вважаються пропусками і заповнюються за методом найближчого сусіда. На рис. 4.11 показано приклад покадрового розпізнавання рівня води на складних зашумлених відеофайлах з низькою роздільною здатністю. На рис. 4.12 показано результат фільтрації даних, приведених на рис. 4.11.



Рисунок 4.11 – Приклад результату покадрового розпізнавання рівня води на даних відеоінспекції труб водовідведення без фільтрації

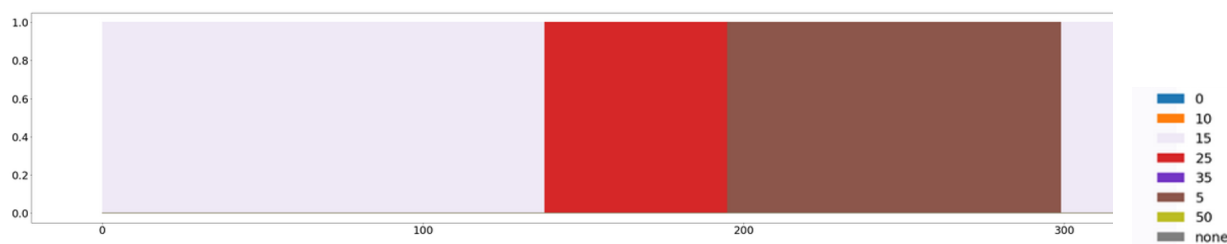


Рисунок 4.12 – Приклад результату покадрового розпізнавання рівня води на даних відеоінспекції труб водовідведення після фільтрації

Аналіз рис. 4.11 та рис. 4.12 показує, що у результатах розпізнавання на неякісних відеозображеннях можливі часті різкі переходи між сусідніми рівнями розпізнавання, які усуваються простими алгоритми фільтрації в рамках однакової відстані або в рамках ковзного вікна 1-2 метри.

Під час оброблення результатів розпізнавання дефектів згортка (фільтрація) відбувається в декілька етапів :

- фільтрація в рамках одного контексту і відстані;
- перекодування контекстів у радіальні координати (рис. 4.8) і фільтрація в рамках окремих радіальних координат (сегментів);
- поглинання структурними дефектами з більшим пріоритетом структурних дефектів з меншим пріоритетом, якщо вони мають однакоvu локалізацію.

Для оцінювання точності алгоритму аналізу відеоінспекцій необхідно обробити значну кількість звітів і порівняти результати з результатами, що виконала людина, або за рахунок підрахунку статистики правок (видалення і додавання записів) на множинів згенерованих системою звітів. В табл. 4.1 показано результати оцінювання ймовірності помилкових рішень за рахунок підрахунку статистики правок згенерованих системою звітів. Правку виконували спеціалісти компанії MOLFAR.AI SP Z O. O. (Гданськ, Польща), що спеціалізуються на аналізі звітів відеоінспекції трубопроводів водовідведення [19]. Для цього було згенеровано і перевірено 100 звітів. Оцінка ймовірності обчислюється шляхом ділення кількості помилкових записів на загальну кількість записів звіту. Кожен звіт містив не менше 50 записів, а кожна інспекція тривала на менше 2 годин відеозйомки.

Аналіз табл. 4.1 показує, що згенеровані звіти за результатами аналізу відеоінспекцій труб, характеризуються найбільшою ймовірністю помилок у випадку полімерного та цегляного матеріалу труб. При цьому під час аналізу полімерних труб найвища ймовірність хибних спрацювань, які видаляє експерт. Проте загалом оцінка повної ймовірності помилкових рішень у звітах про відеоінспекцію труб водовідведення є досить невисокою і прийнятною для практичного використання.

Таким чином, процес формування звіту має багатоетапну структуру. При цьому фільтрація і відновлення даних усувають частину помилок, які допускаються інтелектуальною системою під час аналізу неякісних відеоданих та даних, що знаходяться поза розподілом навчальних вибірок. Оцінка повної ймовірності помилкових рішень під час формування звітів не перевищує 3%, тому точність згенерованих системою звітів прийнятна для практичного використання [103].

Таблиця 4.1 – Оцінка ймовірності помилок у згенерованих системою звітах

Матеріал труби	Оцінка ймовірності пропусків	Оцінка ймовірності хибних спрацювань	Оцінка повної ймовірності помилкових рішень
Полімер	0,01	0,05	0,030
Кераміка та бетон	0,02	0,03	0,025
Цегла	0,03	0,03	0,03
Метал	0,01	0,01	0,01

4.3 Програмний інструментарій та реалізація інтелектуальної системи оцінювання функціонального стану труб водовідведення

Розроблений хмарний сервіс інтелектуального аналізу даних відеоінспекції труб водовідведення складається з клієнтської частини (Frontend Application), яка

надає графічний доступ до платформи та можливість взаємодії з платформою через браузер, та серверної частини (Backend Server), яка обробляє запити від клієнтської сторони. Серверна частина включає в себе серверну логіку з інтелектуальною обробкою даних та логіку роботи з базами і сховищами даних. Бази і сховища даних використовується для зберігання інформації про користувачів та для зберігання завантажених даних. До складу хмарного сервісу входять наступні компоненти (рис. 4.13) [102]:

- Frontend Application, що реалізує клієнтську частину для доступу до веб-сервісу;
- Backend Server, що реалізує серверну частину для обробки запитів користувача;
- Amazon S3 (Amazon Simple Storage Service) – сервіс для зберігання відеоданих;
- S3 Reader – сервіс для сканування сховища Amazon S3 і зчитування даних для подальшого перегляду і інтелектуального оброблення ;
- Elasticsearch – нереляційна база даних для збереження звітів про інспекцію та службової інформації сервісу;
- Kibana – інструмент формування запитів та візуалізації даних Elasticsearch;
- Converter – конвертор відео в уніфікований формат для подальшого аналізу;
- Odometr Recognition – розпізнавач одометру, що нанесений на кадрах відео;
- AI – набір алгоритмів інтелектуального аналізу даних відеоінспекції;
- Result Aggregator – агрегатор результатів інтелектуального оброблення даних для фільтрації викидів і уточнення значень з урахуванням даних одометра та радіальної локалізації дефектів;
- Report Generator – генератор звіту у заданому стандарті.

Сервіс «Frontend Application» це веб додаток, який взаємодіє з користувачем і надає йому доступ до веб інтерфейсу платформи. В цьому інтерфейсі користувачі можуть авторизуватися, завантажувати відео, виконувати пошук, завантажувати відео для оброблення, спостерігати статус оброблення відео, скачувати звіти обраних інспекцій та виконувати зміни особистих налаштувань. На рис. 4.14

показано приклад відображення сторінки з наявними відеоінспекціями та їхніми статусами. Список відеоінспекцій можна фільтрувати за багатьма параметрами, наприклад, місто в якому була інспекція, форма труби, дата інспекції і тому подібне.

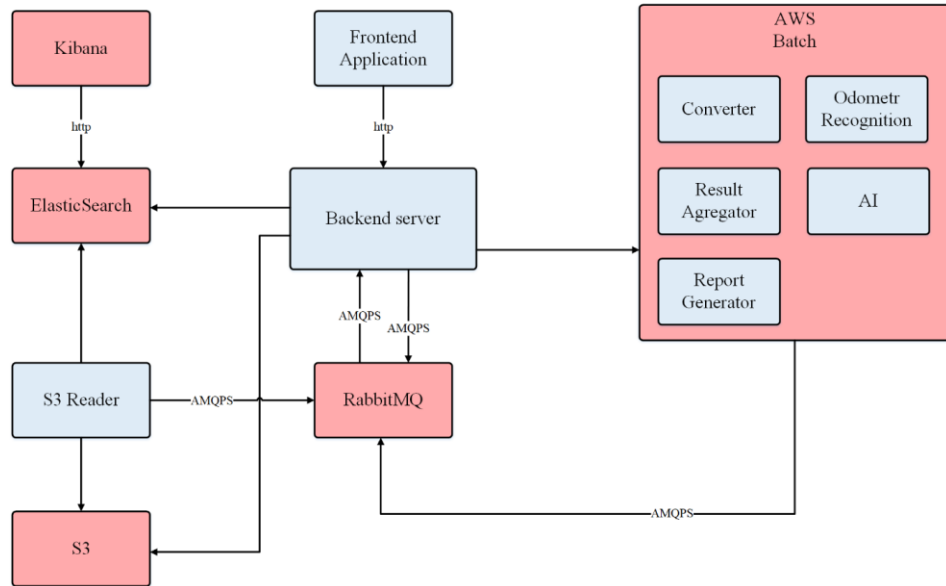


Рисунок 4.13 – Структурна схема хмарного сервісу інтелектуального оброблення даних інспекції труб водовідведення

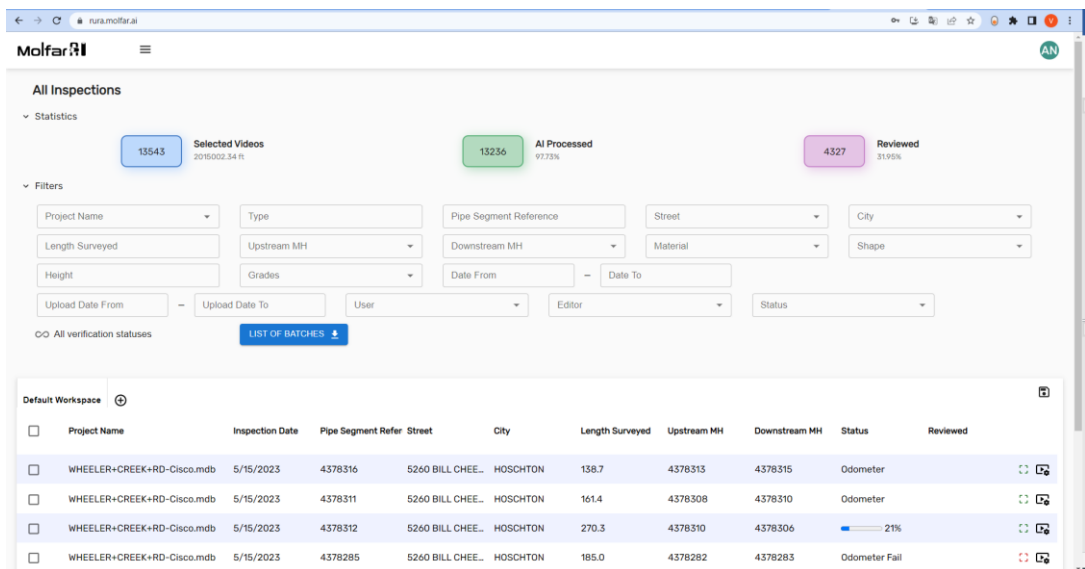


Рисунок 4.14 – Зовнішній вигляд веб-інтерфейсу зі списком відеоінспекцій

На рис. 4.15 показано інтерфейс вже обробленої відеоінспекції. На інтерфейсі наявні такі основні блоки:

- блок 1 з інформацією про інспекцію, де можна побачити, та в разі потреби, відредагувати основну інформацію про інспекцію, наприклад загальну довжину сегменту труби, фактичну довжина даної інспекції, дату та місце проведення інспекції, форму труби, матеріал та багато інших параметрів згідно стандарту;
- блок 2 з відеоплеєром, де відповідно можна переглянути відеоряд з інспекцією;
- блок 3 зі сніпшотами, що відповідають кадрам відео, де були розпізнані ті чи інші коди дефектів;
- блок 4 з записами про коди дефектів у відповідному стандарті, де кожен запис містить інформацію про час, дистанцію коду, назва коду, модифікатор коду, помітки про протяжність коду.

The screenshot shows the Molfar web application interface for video inspection results. The interface is divided into four main sections, each highlighted with a red box and a red number:

- Block 1 (Inspection Info):** A sidebar containing various fields for inspection details, such as Coating Method, Year Constructed, Easting, Elevation, Flow control, GPS accuracy, Location code, Location details, Media label, Northing, Pipe joint length, PO number, Pressure value, Purpose, Reverse setup, and Pipe Use.
- Block 2 (Video):** A central video player showing a live feed of the pipe inspection.
- Block 3 (Snapshots):** A grid of six thumbnail images showing different views of the pipe interior, each labeled with a code (AMH, MWL, CL, SAV, CM, SR).
- Block 4 (Observations):** A table listing inspection observations with columns for Time, Distance, Code, Modifier, Value, Clock At/From, Clock To, Continuous, Dim. 1, Dim. 2, and Remarks.

Time	Distance	Code	Modifier	Value	Clock At/From	Clock To	Continuous	Dim. 1	Dim. 2	Remarks
00:03	0	AMH								S058-504
00:08	0	MWL		5						
01:00	2.3	CL			1					
01:00	2.3	SAV			3	9	S1			
01:07	3.8	CM			10	2	S2			
01:07	3.8	SR			10	12				

Рисунок 4.15 – Вигляд сторінки веб-додатку з результуючими кодами після агрегації результатів

Такі сервіси як Converter, Odometer Recognition, Result Aggregator, AI та Report Generator виконуються як задачі з пакетної обробки. Для пакетної обробки

можуть бути використані Google Cloud's Dataflow, Microsoft's Azure Batch, IBM Spectrum LSF та AWS Batch. Проте AWS Batch має такі переваги [108]:

- динамічне масштабування відповідно до кількості задач;
- забезпечення інформаційної безпеки за рахунок вбудованих в AWS служб та інструментів, таких як IAM для управління доступом, VPC для ізоляції ресурсів, шифрування за допомогою ключів KMS тощо;
- інтеграція з іншими сервісами AWS, такими як Amazon S3, DynamoDB, ECR, ECS, VPC, CloudWatch, і т.д., що дає гнучкість та контроль над додатками та середовищами;
- ефективне управління ресурсами обчислень за рахунок механізмів планування роботи на оптимальних типах EC2-інстансів з огляду на вимоги до роботи і витрати;
- оплата здійснюється тільки за використання AWS ресурсів, необхідних для запуску тієї чи іншої задачі;
- можливість концентрації на написанні та оптимізації додатків, не турбуючись про виконання інфраструктурних завдань, таких як управління кластерами або планування задач;
- підтримка контейнеризації Docker, що дозволяє розробникам упаковувати застосунок разом з усіма його залежностями у стандартний образ, який можна виконувати як незалежний процес на будь-якій платформі, що підтримує Docker [109].

Converter здійснює конвертацію завантаженого відео з інспекцією в необхідний для подальшої обробки формат, роздільну здатність відео та частоту кадрів. Ці перетворення над відео необхідні, оскільки під час конвертації видаляються пошкоджені кадри, якщо вони є, зменшується розмір відео, так як часто оригінальне відео має високу роздільну здатність (FullHD або більше). Конвертація виконується на базі комбінації пайплайнів бібліотеки ffmpeg.

Odometer Recognition знаходить на відео ділянки з текстом, розпізнає текст на цих ділянках, і за декількома шаблонами визначає в якій саме із них розташоване значення одометру. Знати поточне значення одометру на відео також дуже

важливий аспект аналізу відео, оскільки це дає змогу більш точно робити згортку (агрегацію) розпізнаних дефектів, а відповідно підвищувати кінцеву якість отриманих звітів. Для цього було використано популярну бібліотеку «TesseractOCR» [110]. Шляхом експерименту було підбрано найбільш оптимальні параметри роботи бібліотеки, що дозволило зменшити витрату ресурсів без втрат точності розпізнавання одометру.

AI є основним сервісом інтелектуальної обробки відео. Під час роботи цього сервісу використовуються модель для класифікації контексту на відеокадрі, модель аналізу міжкадрових залежностей для уточнення контексту, модель класифікації рівня води на кадрі, модель аналізу міжкадрових залежностей для уточнення рівня води, моделі детекторів дефектів в орієнтації камери спостереження вперед та вбік. Як зазначалось в розділі 2, основною моделлю екстракції ознакового опису спостережень є згорткова мережа загального призначення MobileNet (коефіцієнт ємності даної мережі встановлено на рівні 0,25 та 0,5 а вхідна роздільна здатність 224x224 пікселів). Обрано саме цю модель, оскільки такий екстрактор ознак забезпечує прийнятну для практичного використання точність. При цьому використання більш просунутих моделей для екстракції ознак, наприклад візуальних трансформерів DINOv2 [14] або CLIP [111], потребує помітно більше часу і відповідно коштів на оброблення даних. В табл. 4.2 показано для порівняння показано час оброблення міні-паketу з 128 кадрів відео залежно від архітектури обчислювального пристрою та моделі.

Аналіз табл. 4.2 показує, що моделі з архітектурою DINOv2 (big) та CLIP (big) потребують в декілька разів, або в декілька десятків разів більше часу на оброблення ніж моделі MobileNet($\alpha=0,25$) або MobileNet($\alpha=0,5$). Проте запропоновані методи універсальні і не залежать від міроархітектури екстракторів ознак і результати можуть удосконалюватися за рахунок використання більш просунутих архітектур і мікро-архітектур.

Результати роботи сервісу AI мають вигляд колекцій значень. Одна з колекцій містить значенням рівня води з відміткою часу, а інша колекція містить

класи дефектів з відміткою часу та позиції. Потім ці колекції подаються на мікросервіс Result Aggregator.

Таблиця 4.2 – Час обробки міні-паketу з 128 зображень (в секундах) для різних архітектур мережі та різних обчислювальних пристроїв

Назва архітектури пристрою	Назва MobileNet ($\alpha=0,25$)	Назва MobileNet ($\alpha=0,5$)	DINOv2 big	CLIP big
CPU Intel Xeon (2 ядра)	0,672	1,076	9,553	7,488
GPU A100	0,0071	0,0113	0,347	0,292
GPU P100	0,0106	0,0172	0,486	0,381
GPU T4	0,0401	0,0561	1,214	0,527

Алгоритми Result Aggregator вирішують задачу фільтрації та агрегації результатів розпізнавання сервісів AI та Odometer Recognition. Фільтрація та агрегація результатів відбувається з комплексним урахуванням просторово-часових характеристик дефектів та значень рівня води, а також їх пріоритетності.

Report Generator формує звіт з отриманих даних, сформованим сервісом Result Aggregator, у певному стандарті кодування. Звіт формується у спеціальному файловому форматі бази даних Microsoft Access .mdb де знаходяться дефекти, згідно стандарту, у відповідних до стандарту таблицях. Такий формат є необхідною умовою, щоб мати можливість імпортувати дані із звіту до інших аналітичних програм.

Як база даних та пошукова система для зберігання результатів роботи інтелектуальної системи аналізу даних відеоінспекції труб водовідведення було використано Elasticsearch [112]. Elasticsearch – це відкрите та гнучке програмне забезпечення для зберігання даних, пошуку та аналітики. Надає розподілений, повнотекстовий пошуковий інструмент з HTTP (Hypertext Transfer Protocol) веб-

інтерфейсом і підтримкою безсхемних JSON документів. Elasticsearch дозволяє зберігати, індексувати та аналізувати структуровані та неструктуровані дані в розподілених середовищах з високою доступністю та масштабованістю. До основних характеристик, які вплинули на вибір саме цієї пошукової системи, відносять:

- швидкий пошук інформації за допомогою словосполучень, фраз або регулярних виразів;
- автоматичний розподіл даних між кількома вузлами, що забезпечує відмінну продуктивність та масштабованість;
- можливість створення резервних копій та реплікацій даних, що забезпечує високу доступність та захист від втрати даних;
- можливість роботи з даними без чітких вимог до схеми, що полегшує використання та взаємодію з неструктурованими даними;
- надання RESTful (Representational State Transfer) API (Application Programming Interface) для взаємодії з системою, що дозволяє розробникам легко використовувати Elasticsearch з різними мовами програмування.

Для візуалізації даних з бази використовується інструмент Kibana. Kibana є візуалізаційним та аналітичним інструментом. Він використовується разом з Elasticsearch. Kibana дозволяє візуалізувати та аналізувати дані з Elasticsearch, створюючи графіки, діаграми, карти та інші візуалізації для полегшення роботи з великими обсягами даних. Kibana забезпечує можливість створювати гнучких дашбордів, які дають можливість аналізувати дані в реальному темпі часу, здійснювати пошук та фільтрацію даних, а також співставляти різні дані.

Для зберігання початкових та оброблених системою відеофайлів інспекцій використовується сервіс Amazon S3. Amazon S3 (Simple Storage Service) – це сервіс зберігання в хмарі від Amazon, який дозволяє зберігати, отримувати та захищати будь-який об'єм даних в будь-який час з будь-якої точки глобальної мережі. Переваги Amazon S3 [113]:

- здатність автоматичного масштабування, щоб вміщати великі обсяги даних, від декількох байтів до петабайтів;

- надання розширених функцій безпеки, таких як ACL, політик користувача IAM, шифрування на стороні сервера та клієнта, і багато іншого, що дозволяє чітко розмежувати права доступу до файлів інспекцій та файлів звітів;

- надійність та стійкість за рахунок збереження даних в множинних географічно віддалених локаціях;

- інтеграція з іншими сервісами AWS;

- підтримка декількох класів зберігання даних, що дозволяє обирати компроміси між вартістю та доступністю даних.

Однак вартість S3 може швидко зростати з підвищенням обсягу зберігання даних та трафіку переносу даних, тому необхідно слідкувати за актуальністю тих чи інших даних та переносити їх до іншого класу зберігання для здешевлення утримання. Крім того, хоча S3 і є гнучким, його налаштування та управління не є простою задачею, особливо при використанні розширених функцій безпеки та доступу.

Сервіс S3 Reader виконує роль менеджера завантажень з веб інтерфейсу платформи, формує шляхи та права доступу до файлів для окремих юзерів, відповідно до їхньої ролі, взаємодіє з S3.

Основною компонентою системи є сервіс «Backend Server». Цей сервіс працює на базі відкритої платформи для запуску JavaScript на сервері – Node.js і виконує функцію REST-сервера. Основна особливість Node.js полягає у використанні подійно-орієнтованої та неблокуючої моделі вводу/виводу, яка робить його підходящим для веб-застосунків, що потребують обробки великої кількості одночасних запитів і високої масштабованості. REST – це архітектурний підхід, який використовує прості HTTP-запити для комунікації замість складніших варіантів, таких як CORBA (Common Object Request Broker Architecture), COM + (Component Object Model), RPC (Remote Procedure Call), SOAP (Simple Object Access Protocol). Використання REST означає, що виклики будуть описуватися на основі повідомлень і залежать від стандарту HTTP для опису цих повідомлень. Використання протоколу HTTP означає, що REST – це простий механізм запиту / відповіді. Кожен запит повертає відповідь (рис. 4.12).

Оскільки певні оброблення даних повинно здійснюватися в певній послідовності, а сервіси не мають фіксованої тривалості оброблення, то існує необхідність в забезпеченні асинхронного обміну інформацією між сервісами. Для цього використовують черги. Прикладом реалізації черг є сервіс RabbitMQ. RabbitMQ – це брокер повідомлень з відкритим кодом, який підтримує різні протоколи обміну повідомленнями, включаючи AMQP (Advanced Message Queuing Protocol). RabbitMQ дозволяє програмам і службам надсилати та приймати повідомлення без необхідності перебувати в постійній взаємодії.

З переваг RabbitMQ можна виділити наступні [114]:

- забезпечення функцій постійного зберігання, відправлення підтверджень, автоматичного дзеркала та збору статистики для забезпечення надійності обміну повідомленнями;
- можливості щодо налаштування поведінки обміну повідомленнями, використовуючи маршрутизацію, теми та списки очікування, що дозволяє працювати з різними сценаріями застосування;
- можливості горизонтального та вертикального масштабування;
- багатоплатформенність.

Однак RabbitMQ має певні складності налаштування, оскільки включає багато конфігураційних опцій. Крім того RabbitMQ потребує більше системних ресурсів порівняно з аналогічними рішеннями, такими як Apache Kafka або ActiveMQ. Також RabbitMQ не має достатнього інструментарію для детального та гнучкого моніторингу та аудиту роботи черги.

Вхідні дані зчитуються системою з використанням бібліотеки машинного зору OpenCV і обробляються в форматі масиву NumPy [115]. Проте безпосередній аналіз даних в нейронній мережі відбувається в тензорних масивах бібліотеки Tensorflow [116]. Для простоти написання коду пропонується використовувати високорівневу обгортку Keras. При цьому препроцесінг даних можна здійснювати з використання популярної бібліотеки Scikit-Learn [117].

Таким чином, програмний інструментарій характеризується масштабованістю, високим рівнем інформаційної безпеки та зручністю для пошуку та візуалізації даних і звітів про інспекцію.

Висновки до четвертого розділу

В четвертому розділі розроблено структурні та функціональні схеми, представлено деталі реалізації інформаційного та програмного забезпечення інформаційної технології аналізу даних відеоінспекції водовідведення, що функціонує за умов інформаційних та ресурсних обмежень. В результаті проведених досліджень і програмної реалізації запропонованих алгоритмів отримано наступні результати:

1. Розроблено узагальнену та деталізовану функціональну модель процесу оптимізації параметрів системи класифікаційного аналізу даних в рамках запропонованої інформаційної технології. Перевагою запропонованого рішення є ефективне використання розмічених і нерозмічених даних для поетапної ініціалізації вагових коефіцієнтів моделі екстрактора ознак та побудові збагачених інформацією цільових міток у вигляді двійкових еталонних векторів (прототипів класів), що використовуються для точної настройки системи. При цьому калібрування порогів спрацювання здійснюється в процесі оптимізації за інформаційним критерієм меж гіперсферичних контейнерів класів розпізнавання. Даний підхід знижує вимоги до обсягу розмічених навчальних даних та підвищує узагальнюючу здатність системи аналізу даних відеоінспекції.

2. Розроблено узагальнену та деталізовану функціональну модель процесу аналізу даних відеоінспекції для формування звітів в рамках запропонованої інформаційної технології. Перевагою запропонованого рішення є декомпозиція задачі аналізу даних на підзадачі класифікації контексту спостереження, рівня води та дефектів у визначених ділянках зображення залежно від контексту, що підвищує прозорість системи і спрощує збір розмічених навчальних даних для настройки даної системи. При цьому аналіз міжкадрових залежностей під час екстракції ознак

та використання агрегації результатів за значеннями одометра і пріоритетом дефектів разом з фільтрацією викидів і пропусків дозволяють підвищити достовірність результату за умов ресурсних та інформаційних обмежень.

3. Запропоновано структуру і програмну реалізацію хмарного сервісу для аналізу даних відеоінспекції з автоматичним формуванням звітів. Перевагою даного рішення є можливість асинхронної обробки множини відео за рахунок черги повідомлень та пакетного оброблення задач. Функція веб-інтерфейсу і веб-сервера полягає у завантаженні даних, відправленні даних до пакетного оброблення і регламентація етапів оброблення. При цьому контейнеризація задач оброблення даних спрощує їх розроблення і підтримання.

Основні наукові результати, наведені у четвертому розділі, опубліковано у працях автора [19], [78], [81], [102], [103], [100].

ВИСНОВКИ

У дисертаційній роботі виконано важливе наукове завдання — підвищення ефективності відеоінспекції труб водовідведення шляхом розроблення інформаційної технології інтелектуального аналізу даних.

1. У дисертаційній роботі визначено актуальність застосування інформаційних технологій у галузі інспекції труб водовідведення. На основі проведеного аналізу методів та підходів до інспекції труб водовідведення встановлено, що основним джерелом інформації є відеокамери, а основна тенденція розвитку сучасних технологій аналізу даних відеоінспекції пов'язана з впровадженням технологій машинного зору та штучного інтелекту. За результатами аналізу процесу відеоінспекції труб водовідведення обґрунтовано потребу у створенні нової інтелектуальної інформаційної технології оцінювання функціонального стану труб водовідведення шляхом використання ідей і методів композиції простих моделей, контрастного навчання, завадозахищеного кодування та інформаційно-екстремальної інтелектуальної технології.

2. Уперше розроблено метод навчання ієрархічного екстрактора ознак, що на відміну від відомих, забезпечує підвищення функціональної ефективності вирішувальних правил за умов зашумленості та незбалансованості навчальних даних шляхом поєднання принципів переносу знань, сіамських мереж, контрастного самонавчання та самокоректуючих двійкових кодів. Запропонований метод машинного навчання забезпечує точність 98% для класифікаційного аналізу дефектів у контексті спостереження з орієнтацією камери вперед вздовж труби, що перевищує на 11% результат традиційного підходу до навчання з кінця-в-кінець з повнозв'язним шаром класифікації. Оскільки в контексті спостереження, що відповідає орієнтації камери на бічну стінку, присутня помітна незбалансованість даних, то ефективність класифікації дефектів оцінюється F1-метрикою і становить $F1=0,977$, що на 5% більше ніж в рамках традиційного підходу. Отримана точність класифікації дефектів в рамках запропонованого підходу перевищує на 6%

результати отримані іншими авторами [25], і дозволяє отримати більшу точність звітів про функціональний стан труб водовідведення.

3. Удосконалено моделі екстракції ознакового опису спостережень шляхом поєднання локальної і просторово-часової візуальної контекстної інформації, що дозволяє підвищити інформативність ознакового опису відеоспостережень інспекції трубопроводу водовідведення за умов неповної визначеності. Встановлено перевагу використання 1D згорткової моделі порівняно з LSTM та GRU мережами для аналізу міжкадрових залежностей. Мікроусереднене значення F1-міри навченого в рамках запропонованого підходу екстрактора ознак для розпізнавання контексту спостережень з алфавіту 11 класів становить 0,92, що перевищує на 5% результат для моделі без врахування міжкадрових залежностей. Для алфавіту класів, що описує рівні води від 0% до 55% з кроком в 5% мікроусереднене значення F1-міри в рамках запропонованого підходу становить 0,879 і перевищує на 9% значення для моделі без врахування міжкадрових залежностей.

4. Набули подальшого розвитку моделі і алгоритми синтезу вирішувальних правил для класифікаційного аналізу спостережень, які основані на композиції моделей взаємодоповнюючих задач та оптимізації параметрів класифікаційних вирішувальних правил та гіперпараметрів екстрактора ознак за інформаційним критерієм, що є функцією результату перевірки статистичних гіпотез. Тобто моделі з інформаційно-екстремальними вирішувальними правилами для розпізнавання контексту, рівня води та дефектів komponуються для оцінювання функціонального стану труби. За результатами перевірки 100 звітів про інспекцію, що формують за результатами агрегації результатів розпізнавання вздовж труби, отримано оцінку повної ймовірності помилкових рішень під час формування звітів, яка не перевищує 3%, що свідчить про прийнятний для практичного використання рівень точності згенерованих системою звітів.

5. У результаті виконаного дослідження розроблено інформаційне та програмне забезпечення інтелектуальної системи, яка розгортається у хмарному середовищі, для автоматичного формування звіту про функціональний стан труб

водовідведення за даними відеоінспекції.

6. Одержані наукові результати досліджень у вигляді інформаційного та програмного забезпечення впроваджено під час підготовки персоналу та проведення інспекції каналізаційної мережі в КП "Міськводоканал" СМР (м. Суми), під час розроблення інтелектуальної системи відеомоніторингу інфраструктурних об'єктів в Товаристві з обмеженою відповідальністю "РОЗУМНІ ТЕХНОЛОГІЇ «ТИТУЛ»" (м. Суми), під час розроблення програмного забезпечення модуля машинного зору для дистанційно-керованої мобільної платформи в Товаристві з обмеженою відповідальністю "НОРД ТРЕК" (м. Шостка, Сумська область), під час розроблення системи класифікаційного аналізу даних відеоінспекції трубопроводу водовідведення в компанії Molfar.AI sp. z o.o., (місто Гданськ, Польща), а також у навчальному процесі кафедри комп'ютерних наук Сумського державного університету під час викладання дисципліни "Introduction to Data Science".

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. A. AlShami *et al.*, “Performance Assessment of Sewer Networks under Different Blockage Situations Using Internet-of-Things-Based Technologies,” *Sustainability*, vol. 14, no. 21, p. 14036, Oct. 2022, doi: 10.3390/su142114036.
2. S. Moradi, T. Zayed, and F. Golkhoo, “Review on Computer Aided Sewer Pipeline Defect Detection and Condition Assessment,” *Infrastructures*, vol. 4, no. 1, p. 10, Mar. 2019, doi: 10.3390/infrastructures4010010.
3. *Infrastructure Report Card*, Wastewater, American Society of Civil Engineers (ASCE). – 2017, pp.31-47. Available at: <https://www.infrastructurereportcard.org/wp-content/uploads/2017/01/Wastewater-Final.pdf>. (accessed on 5 May 2021).
4. J. C. Y. Cheng and M. Wang, “Automated detection of sewer pipe defects in closed-circuit television images using deep learning techniques,” *Automation in Construction*, vol. 95, pp. 155–171, Nov. 2018, doi: 10.1016/j.autcon.2018.08.006.
5. P. Palanisamy *et al.*, “Drain Structural Defect Detection and Mapping Using AI-Enabled Reconfigurable Robot Raptor and IoRT Framework,” *Sensors*, vol. 21, no. 21, p. 7287, Nov. 2021, doi: 10.3390/s21217287.
6. H. Zhou, C. Xu, X. Tang, S. Wang, and Z. Zhang, “A Review of Vision-Laser-Based Civil Infrastructure Inspection and Monitoring,” *Sensors*, vol. 22, no. 15, p. 5882, Aug. 2022, doi: 10.3390/s22155882.
7. Q. Ping *et al.*, “Pipeline In-Line Inspection Method, Instrumentation and Data Management,” *Sensors*, vol. 21, no. 11, p. 3862, Jun. 2021, doi: 10.3390/s21113862.
8. F. Huang, N. Wang, H. Fang, H. Liu, and G. Pang, “Research on 3D Defect Information Management of Drainage Pipeline Based on BIM,” *Buildings*, vol. 12, no. 2, p. 228, Feb. 2022, doi: 10.3390/buildings12020228.
9. Y. Li, H. Wang, L. M. Dang, H.-K. Song, and H. Moon, “Vision-Based Defect Inspection and Condition Assessment for Sewer Pipes: A Comprehensive Survey,” *Sensors*, vol. 22, no. 7, p. 2722, Apr. 2022, doi: 10.3390/s22072722.
10. X.-J. Zhang, P. Zhao, Q. Hu, H. Wang, M. Ai, and J. Li, “A 3D Reconstruction Pipeline of Urban Drainage Pipes Based on Multiview Image Matching Using Low-Cost

Panoramic Video Cameras,” *Water*, vol. 11, no. 10, p. 2101, Oct. 2019, doi: 10.3390/w11102101.

11. J. F. C. Sham, W. W. L. Lai, B. P. Chan, and C. G. Koh, “Imaging and condition diagnosis of underground sewer liners via active and passive infrared thermography: A case study in Singapore,” *Tunnelling and Underground Space Technology*, vol. 84, pp. 440–450, Feb. 2019, doi: 10.1016/j.tust.2018.11.013.

12. F. Hou, X. Rui, X. Fan, and H. Zhang, “Review of GPR Activities in Civil Infrastructures: Data Analysis and Applications,” *Remote Sensing*, vol. 14, no. 23, p. 5972, Nov. 2022, doi: 10.3390/rs14235972.

13. T.-Y. Chuang and C. K. Sung, “Learning and SLAM Based Decision Support Platform for Sewer Inspection,” *Remote Sensing*, vol. 12, no. 6, p. 968, Mar. 2020, doi: 10.3390/rs12060968.

14. C. H. Bahnsen, A. Johansen, M. P. Philipsen, J. W. Henriksen, K. Nasrollahi, and T. B. Moeslund, “3D Sensors for Sewer Inspection: A Quantitative Review and Analysis,” *Sensors*, vol. 21, no. 7, p. 2553, Apr. 2021, doi: 10.3390/s21072553.

15. D. Alejo, C. Marques, F. Caballero, P. Alvito, and L. Merino, *SIAR: an autonomous ground robot for sewer inspection*. 2022. doi: 10.17979/spudc.9788497498081.1198.

16. H. Plihal *et al.*, “A novel method for rapid inspection of sewer networks: combining acoustic and optical means,” *Urban Water Journal*, vol. 13, no. 1, pp. 3–14, Sep. 2015, doi: 10.1080/1573062x.2015.1076857.

17. WRc Knowledge Store, “Manual of Sewer Condition Classification 5th Edition,” 2013. *WRc Knowledge Store*. <https://wrcknowledgestore.co.uk/products/manual-of-sewer-condition-classification-5th-edition> (accessed Feb. 02, 2022).

18. *Report Card for America’s Infrastructure*, Wastewater, Available online: <https://infrastructurereportcard.org/wp-content/uploads/2020/12/Wastewater-2021.pdf> (accessed Feb. 20, 2022).

19. V. V. Moskalenko, M. O. Zaretskyi, A. S. Moskalenko, and V. Lysyuk, “Sewer Pipe Defects Classification Based on Deep Convolutional Network with Information-

Extreme Error-Correction Decision Rules,” in *Springer eBooks*, 2020, pp. 253–263. doi: 10.1007/978-3-030-61656-4_16.

20. WinCan, “WinCan - Smarter Sewer Inspection | AI.” <https://blog.wincan.com/topic/ai> (accessed Jan. 25, 2022).

21. “AutoCode | SewerAI.” <https://www.sewerai.com/products/autocode> (accessed Jan. 21, 2022).

22. В.В. Москаленко, А. А. Москаленко, А.Г. Коробов, М. О. Зарецький, та В.А. Семашко, “Модель та алгоритм навчання системи детектування малорозмірних об’єктів для малогабаритних безпілотних літальних апаратів,” *Радіоелектронні І Комп’ютерні Системи*, № 4, С. 41–52, 2018, doi: 10.32620/reks.2018.4.04.

23. В. В. Москаленко, А. С. Москаленко, М. О. Зарецький, та А. Г. Коробов, “Модель та алгоритм навчання системи автономної навігації для малогабаритного безпілотного апарату ,” *Біоніка інтелекту*, № 91, С. 16–24, 2018.

24. В. В. Москаленко, М. О. Зарецький, та А. А. Москаленко, “Модель та метод навчання системи класифікації рухомих об’єктів для малогабаритного безпілотного апарату,” *Радіоелектронні І Комп’ютерні Системи*, №. 2, С. 108–117, 2019, doi: 10.32620/reks.2019.2.10.

25. N. M. Syahrian, P. Risma, and T. Dewi, “Vision-Based Pipe Monitoring Robot for Crack Detection Using Canny Edge Detection Method as an Image Processing Technique,” *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, pp. 243–250, Sep. 2017, doi: 10.22219/kinetik.v2i4.243.

26. J. B. Haurum and T. B. Moeslund, “A Survey on Image-Based Automation of CCTV and SSET Sewer Inspections,” *Automation in Construction*, vol. 111, p. 103061, Mar. 2020, doi: 10.1016/j.autcon.2019.103061.

27. X. Fang *et al.*, “Sewer Pipeline Fault Identification Using Anomaly Detection Algorithms on Video Sequences,” *IEEE Access*, vol. 8, pp. 39574–39586, Jan. 2020, doi: 10.1109/access.2020.2975887.

28. J. Myrans, R. M. Everson, and Z. Kapelan, “Automated detection of fault types in CCTV sewer surveys,” *Journal of Hydroinformatics*, vol. 21, no. 1, pp. 153–163, Oct. 2018, doi: 10.2166/hydro.2018.073.

29. F. Panella, J. Boehm, Y.-L. Loo, A. Kaushik, and D. Gonzalez, “Deep Learning and Image Processing for Automated Crack Detection and Defect Measurement In Underground Structures,” *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLII–2, pp. 829–835, May 2018, doi: 10.5194/isprs-archives-xlii-2-829-2018.

30. D. Li, Q. Xie, Z. Yu, Q. Wu, J. Zhou, and J. Wang, “Sewer pipe defect detection via deep learning with local and global feature fusion,” *Automation in Construction*, vol. 129, p. 103823, Sep. 2021, doi: 10.1016/j.autcon.2021.103823.

31. D. Li, A. Cong, and S. Guo, “Sewer damage detection from imbalanced CCTV inspection data using deep convolutional neural networks with hierarchical classification,” *Automation in Construction*, vol. 101, pp. 199–208, May 2019, doi: 10.1016/j.autcon.2019.01.017.

32. Y. Bengio, A. Courville, and P. Vincent, “Representation Learning: A Review and New Perspectives,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798–1828, Aug. 2013, doi: 10.1109/tpami.2013.50.

33. Q. Li, B. Li, J. M. Garibaldi, and G. Qiu, “Clustering-Based Representation Learning through Output Translation and Its Application to Remote-Sensing Images,” *Remote Sensing*, vol. 14, no. 14, p. 3361, Jul. 2022, doi: 10.3390/rs14143361.

34. K. Pasupa and W. Sunhem, “A comparison between shallow and deep architecture classifiers on small dataset,” *2016 8th International Conference on Information Technology and Electrical Engineering (ICITEE)*, Oct. 2016, doi: <https://doi.org/10.1109/iciteed.2016.7863293>.

35 J. T. Springenberg and M. Riedmiller, “Learning Temporal Coherent Features through Life-Time Sparsity,” in *Lecture Notes in Computer Science*, Springer Science+Business Media, 2012. doi: 10.1007/978-3-642-34475-6_42.

36. X. Chen, C. Gong, C. Ma, X. Huang, and J. Yang, “Privileged Semi-Supervised Learning,” *2018 25th IEEE International Conference on Image Processing (ICIP)*. doi: 10.1109/icip.2018.8451098.
37. Y. Zhang and Q. Yang, “A Survey on Multi-Task Learning,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 12, pp. 5586–5609, Dec. 2022, doi: 10.1109/tkde.2021.3070203.
38. A. M. Saxe *et al.*, “On the information bottleneck theory of deep learning,” *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2019, no. 12, p. 124020, Dec. 2019, doi: 10.1088/1742-5468/ab3985.
39. Y.-H. A. Lee, W.-J. Kim, W.-P. Park, and S. Choi, “Discrete Infomax Codes for Supervised Representation Learning,” *Entropy*, vol. 24, no. 4, p. 501, Apr. 2022, doi: 10.3390/e24040501.
40. A. G. Liu, S. Jin, C.-I. Lai, A. Rouditchenko, A. Oliva, and J. Glass, *Cross-Modal Discrete Representation Learning*. 2022. doi: 10.18653/v1/2022.acl-long.215.
41. H. J. Patel and H. Mewada, “Dictionary Properties for Sparse Representation: Implementation and Analysis,” *Journal of Artificial Intelligence*, vol. 11, no. 1, pp. 1–8, Dec. 2017, doi: 10.3923/jai.2018.1.8.
42. H. Youn, S. Kwon, H.-H. Lee, J. Kim, S.-N. Hong, and D.-J. Shin, “Construction of Error Correcting Output Codes for Robust Deep Neural Networks Based on Label Grouping Scheme,” *2021 7th IEEE International Conference on Network Intelligence and Digital Content (IC-NIDC)*, pp. 17–19, Nov. 2021. doi: 10.1109/icnidc54101.2021.9660486.
43. A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, “A survey of the recent architectures of deep convolutional neural networks,” *Artificial Intelligence Review*, vol. 53, no. 8, pp. 5455–5516, Apr. 2020, doi: 10.1007/s10462-020-09825-6.
44. C. Subakan, M. Ravanelli, S. Cornell, M. Bronzi, and J. Zhong, “Attention Is All You Need In Speech Separation,” *2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. doi: 10.1109/icassp39728.2021.9413901.
45. T. Lin, Y. Wang, X.-Y. Liu, and X. Qiu, “A survey of transformers,” *AI Open*, vol. 3, pp. 111–132, Jan. 2022, doi: 10.1016/j.aiopen.2022.10.001.

46. S. Khan, M. Naseer, M. Hayat, S. W. Zamir, F. S. Khan, and M. Shah, “Transformers in Vision: A Survey,” *ACM Computing Surveys*, vol. 54, no. 10s, pp. 1–41, Jan. 2022, doi: 10.1145/3505244.
47. V. V. Moskalenko, V.S. Kharchenko, A.S. Moskalenko, and S.O. Petrov, “Model and Training Method of the Resilient Image Classifier Considering Faults, Concept Drift, and Adversarial Attacks,” *Algorithms*, vol. 15, no. 10, p. 384, Oct. 2022, doi: 10.3390/a15100384.
48. X. Yin, Y. Chen, A. Bouferguene, H. Zaman, M. Al-Hussein, and L. Kurach, “A deep learning-based framework for an automated defect detection system for sewer pipes,” *Automation in Construction*, vol. 109, p. 102967, Jan. 2020, doi: 10.1016/j.autcon.2019.102967.
49. E.-J. Ong, S. A. Husain, M. Bober-Irizar, and M. Bober, “Deep Architectures and Ensembles for Semantic Video Classification,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 12, pp. 3568–3582, Dec. 2019, doi: 10.1109/tcsvt.2018.2881842.
50. Xi, R., Hou, M., Fu, M., Qu, H., & Liu, D. “Deep Dilated Convolution on Multimodality Time Series for Human Activity Recognition,” *IEEE International Joint Conference on Neural Networks (IJCNN)* 2018, <https://doi.org/10.1109/ijcnn.2018.8489540>
51. X.-B. Jin, X.-H. Yu, X. Wang, Y. Bai, T. Su, and J.-L. Kong, “Prediction for Time Series with CNN and LSTM,” in *Lecture notes in electrical engineering*, Springer Science+Business Media, 2019, pp. 631–641. doi: 10.1007/978-981-15-0474-7_59.
52. Y. W. Hong, M. Mundt, S. Park, Y. Uh, and H. Byun, “Return of the normal distribution: Flexible deep continual learning with variational auto-encoders,” *Neural Networks*, vol. 154, pp. 397–412, Oct. 2022, doi: 10.1016/j.neunet.2022.07.016.
53. P. H. Le-Khac, G. Healy, and A. F. Smeaton, “Contrastive Representation Learning: A Framework and Review,” *IEEE Access*, vol. 8, pp. 193907–193934, Jan. 2020, doi: 10.1109/access.2020.3031549.

54. A. Khan, S. A. AlBarri, and M. F. Manzoor, “Contrastive Self-Supervised Learning: A Survey on Different Architectures,” *In 2022 2nd International Conference on Artificial Intelligence (ICAI)*. 2022 2nd International Conference on Artificial Intelligence (ICAI). *IEEE*, 2022. doi: 10.1109/icai55435.2022.9773725.

55. S. A. R. Zaidi, M. Ansari, A. Aslam, N. Kanwal, M. N. Asghar, and B. R. Lee, “A survey of modern deep learning based object detection models,” *Digital Signal Processing*, vol. 126, p. 103514, Jun. 2022, doi: 10.1016/j.dsp.2022.103514.

56. S. R. Prakash and P. B. Singh, “Object detection through region proposal based techniques,” *Materials Today: Proceedings*, vol. 46, pp. 3997–4002, Jan. 2021, doi: 10.1016/j.matpr.2021.02.533.

57. J. Uijlings, K. E. A. Van De Sande, T. Gevers, and A. W. M. Smeulders, “Selective Search for Object Recognition,” *International Journal of Computer Vision*, vol. 104, no. 2, pp. 154–171, Apr. 2013, doi: 10.1007/s11263-013-0620-5.

58. L. Zhu, F. Lee, J. Cai, H. Yu, and Q. Y. Chen, “An improved feature pyramid network for object detection,” *Neurocomputing*, vol. 483, pp. 127–139, Apr. 2022, doi: 10.1016/j.neucom.2022.02.016.

59. N. Ravi, S. Naqvi, and M. El-Sharkawy, “BIOU: An Improved Bounding Box Regression for Object Detection,” *Journal of Low Power Electronics and Applications*, vol. 12, no. 4, p. 51, Sep. 2022, doi: 10.3390/jlpea12040051.

60. N. Ravi and M. El-Sharkawy, “Improved Single Shot Detector with Enhanced Hard Negative Mining Approach,” *In 2022 International Conference on Advanced Computer Science and Information Systems (ICACISIS)*, 2022 doi: 10.1109/icacsis56558.2022.9923534.

61. N. M. Krishna, R. Y. Reddy, M. S. C. Reddy, K. P. Madhav, and G. Sudham, “Object Detection and Tracking Using Yolo,” *In 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA)*, 2021. doi: 10.1109/icirca51532.2021.9544598.

62. K. Oksuz, B. C. Cam, S. Kalkan, and E. Akbas, “Imbalance Problems in Object Detection: A Review,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 10, pp. 3388–3415, Oct. 2021, doi: 10.1109/tpami.2020.2981890.

63. J. Li, Y. Guo, S. Lao, Y. Wu, L. Bai, and Y. Wei, “Towards a high robust neural network via feature matching,” *International Journal of Multimedia Information Retrieval*, vol. 10, no. 4, pp. 227–237, Oct. 2021, doi: 10.1007/s13735-021-00219-0.

64. M. Kaya and H. S. Bilge, “Deep Metric Learning: A Survey,” *Symmetry*, vol. 11, no. 9, p. 1066, Aug. 2019, doi: 10.3390/sym11091066.

65. Z. Gong, P. Zhong, and W. Hu, “Diversity in Machine Learning,” *IEEE Access*, vol. 7, pp. 64323–64350, Jan. 2019, doi: 10.1109/access.2019.2917620.

66. В. В. Москаленко, *Моделі і методи інтелектуального аналізу багатовимірних даних за умов апріорної невизначеності*. Суми, Україна: Сумський державний університет, 2020. – 184 с.

67. M. Ali, Y. H. Sahin, S. Ozogur-Akyuz, G. Unal, and B. Ç. Otar, “Tuning Accuracy-Diversity Trade-off in Neural Network Ensemble via Novel Entropy Loss Function,” *13th International Conference on Electrical and Electronics Engineering (ELECO)*, 2021. doi: 10.23919/eleco54474.2021.9677845.

68. G. Verma and A. Swami, *Error Correcting Output Codes Improve Probability Estimation and Adversarial Robustness of Deep Neural Networks*, vol. 32. 2019, pp. 8643–8653. Available: <http://papers.nips.cc/paper/9070-error-correcting-output-codes-improve-probability-estimation-and-adversarial-robustness-of-deep-neural-networks.pdf>

69. А. С. Довбиш, І. В. Шелехов, Ю. О. Хібовська, та О. І. Матяш, “Інформаційно-аналітична система оцінювання відповідності сучасним вимогам навчального контенту спеціальності кібербезпека,” *Радіоелектронні І Комп’ютерні Системи*, № 1, С. 70–80, 2021, doi: 10.32620/reks.2021.1.06.

70. В. В. Москаленко, М. О. Зарецький, А. С. Москаленко, А. Г. Коробов, та Я. Ю. Ковальський, “Багатоетапний метод глибинного навчання з попереднім самонавчанням для класифікаційного аналізу дефектів стічних труб,” *Радіоелектронні І Комп’ютерні Системи*, № 4, С. 71–81, 2021, doi: 10.32620/reks.2021.4.06.

71. J. V. Haurum, C. H. Bahnsen, M. Pedersen, and T. B. Moeslund, “Water Level Estimation in Sewer Pipes Using Deep Convolutional Neural Networks,” *Water*, vol. 12, no. 12, p. 3412, Dec. 2020, doi: 10.3390/w12123412.

72. J. Bourcier, T. Floquet, G. Dashyan, T. Ceillier, K. Alahari, and J. Chaussoot, “Self-Supervised Pretraining on Satellite Imagery: a Case Study on Label-Efficient Vehicle Detection,” *arXiv (Cornell University)*, Oct. 2022, doi: 10.48550/arxiv.2210.11815.

73. F. Küppers, A. Haselhoff, J. Kronenberger, and J. Schneider, “Confidence Calibration for Object Detection and Segmentation,” in *Springer eBooks*, 2022, pp. 225–250. doi: 10.1007/978-3-031-01233-4_8.

74. Q. Wang, Y. Ma, K. Zhao, and Y. Tian, “A Comprehensive Survey of Loss Functions in Machine Learning,” *Annals of Data Science*, vol. 9, no. 2, pp. 187–212, Apr. 2020, doi: 10.1007/s40745-020-00253-5.

75. B. C. Geiger and G. Kubin, “Information Bottleneck: Theory and Applications in Deep Learning,” *Entropy*, vol. 22, no. 12, p. 1408, Dec. 2020, doi: 10.3390/e22121408.

76. K. Bhardwaj, C.-Y. Lin, A. L. Sartor, and R. Marculescu, “Memory- and Communication-Aware Model Compression for Distributed Deep Learning Inference on IoT,” *ACM Transactions in Embedded Computing Systems*, vol. 18, no. 5s, pp. 1–22, Oct. 2019, doi: 10.1145/3358205.

77. Y. Gao *et al.*, “Estimating GPU memory consumption of deep learning models,” *28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering. ESEC/FSE*, 2020. doi: 10.1145/3368089.3417050.

78. В. В. Москаленко, М. О. Зарецький, Я. Ю. Ковальський, та С. С. Мартиненко, “Модель і метод навчання класифікатора контекстів спостереження на зображеннях відеоінспекції стічних труб,” *Радіоелектронні І Комп’ютерні Системи*, № 3, С. 59–66, 2020, doi: 10.32620/reks.2020.3.06.

79. В. В. Москаленко, М. О. Зарецький, А. С. Москаленко, А. Г. Коробов, та Я. Ю. Ковальський, “Багатоетапний метод глибинного навчання з попереднім самонавчанням для класифікаційного аналізу дефектів стічних труб,” *Radioelectronic and Computer Systems*, № 4, С. 71–81, 2021, doi: 10.32620/reks.2021.4.06.

80. V. V. Moskalenko, M. O. Zaretsky, A. S. Moskalenko, A. O. Panych, and V. V. Lysyuk, "A Model and Training Method for Context Classification in Cctv Sewer Inspection Video Frames," *Radio Electronics, Computer Science, Control*, no. 3, pp. 97–108, Oct. 2021, doi: <https://doi.org/10.15588/1607-3274-2021-3-9>.

81. В. В. Москаленко, М. О. Зарецький, А. Г. Коробов, Я. Ю. Ковальський, А. Ф. Шаєхов, В. А. Семашко, та ін., "Модель та метод навчання для класифікаційного аналізу рівня води в стічних трубах за даними відео інспекції," *Радіоелектронні І Комп'ютерні Системи*, № 2, С. 4–15, 2021, doi: [10.32620/reks.2021.2.01](https://doi.org/10.32620/reks.2021.2.01).

82. Q. Dai, J. Li, J. Wang, and Y.-G. Jiang, "Binary Optimized Hashing," *24th ACM international conference on Multimedia*, 2016. doi: [10.1145/2964284.2964331](https://doi.org/10.1145/2964284.2964331).

83. A. M. Iiyasu, A. M. Mayet, R. Hanus, F. E. A. El-Samie, and A. Salama, "Employing GMDH-Type Neural Network and Signal Frequency Feature Extraction Approaches for Detection of Scale Thickness inside Oil Pipelines," *Energies*, vol. 15, no. 12, p. 4500, Jun. 2022, doi: [10.3390/en15124500](https://doi.org/10.3390/en15124500).

84. V. V. Moskalenko, A. S. Moskalenko, A. G. Korobov, and M. O. Zaretsky, "A Model and Training Algorithm of Small-Sized Object Detection System for a Compact Aerial Drone," *Radio Electronics, Computer Science, Control*, vol. 0, no. 1, Apr. 2019, doi: <https://doi.org/10.15588/1607-3274-2019-1-11>.

85. В. В. Москаленко, А. С. Москаленко та М. О. Зарецький, "Модель і алгоритм навчання детектора шкідливого трафіку на основі модифікації зростаючого нейронного газу," *Радіоелектронні І Комп'ютерні Системи*, № 3, С. 11–19, 2018, doi: [10.32620/reks.2018.3.02](https://doi.org/10.32620/reks.2018.3.02).

86. V. V. Moskalenko, A.S. Moskalenko, M.O. Zaretskyi, and V. Lysyuk, "Deep Feature Extractor with Information-Extreme Decision Rules for Visual Classification of Sewer Pipe Defects and its Training Method," *Data Stream Mining & Processing*, pp. 191–195, 2022 doi: [10.1109/dsmp47368.2020.9204083](https://doi.org/10.1109/dsmp47368.2020.9204083).

87. I. Shafi, A. Mazahir, A. Fatima, and I. Ashraf, "Internal defects detection and classification in hollow cylindrical surfaces using single shot detection and MobileNet," *Measurement*, vol. 202, p. 111836, Oct. 2022, doi: [10.1016/j.measurement.2022.111836](https://doi.org/10.1016/j.measurement.2022.111836).

88. B. Lim and S. Zohren, “Time Series Forecasting With Deep Learning: A Survey,” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 379, no. 2194, p. 20200209, Apr. 2021, doi: <https://doi.org/10.1098/rsta.2020.0209>.

89. V. V. Moskalenko, A. S. Moskalenko, A. G. Korobov, and M. O. Zaretsky, “Image Classifier Resilient to Adversarial Attacks, Fault Injections and Concept Drift – Model Architecture and Training Algorithm,” *Radio Electronics, Computer Science, Control*, no. 3, p. 86, Oct. 2022, doi: 10.15588/1607-3274-2022-3-9.

90. X. Liu *et al.*, “Self-supervised Learning: Generative or Contrastive,” *IEEE Transactions on Knowledge and Data Engineering*, p. 1, Jan. 2021, doi: 10.1109/tkde.2021.3090866.

91. T. Konkle and G. A. Alvarez, “A self-supervised domain-general learning framework for human ventral stream representation,” *Nature Communications*, vol. 13, no. 1, Jan. 2022, doi: 10.1038/s41467-022-28091-4.

92. H. Kariuki, S. Mwalili, A. Waititu. Dimensionality Reduction of Data with Neighbourhood Components Analysis, *International Journal of Data Science and Analysis*. Volume 8, Issue 3, June 2022 , pp. 72-81. doi: 10.11648/j.ijdsa.20220803.11

93. V. V. Moskalenko, A. S. Moskalenko, and M. O. Zaretskyi, “Model and training method of moving object classification system for a compact drone,” *14th International Conference on Computer Sciences and Information Technologies (CSIT)*, 2019, pp. 56–61. doi: 10.1109/stc-csit.2019.8929795.

94. . M. Zaretskyi V. Moskalenko, A. Moskalenko, and A. Shaiekhov, “Multi-Layer Model and Training Method for Information-Extreme Malware Traffic Detector,” *The Third International Workshop on Computer Modeling and Intelligent Systems* , April 27-May 1, 2020, pp. 288-299 Available at: <https://ceur-ws.org/Vol-2608/paper22.pdf>. (accessed on May 1 2021).

95. D. M. Belete and M. D. Huchaiah, “Grid search in hyperparameter optimization of machine learning models for prediction of HIV/AIDS test results,” *International Journal of Computers and Applications*, vol. 44, no. 9, pp. 875–886, Sep. 2021, doi: 10.1080/1206212x.2021.1974663.

96. H. E. Romeijn, “Random Search Methods,” in *Springer eBooks*, 2008, pp. 3245–3251. doi: 10.1007/978-0-387-74759-0_556.
97. Y.-S. Lee and P.-J. Moon, “Analysis of Open-Source Hyperparameter Optimization Software Trends,” *International Journal of Advanced Culture Technology*, vol. 7, no. 4, pp. 56–62, Dec. 2019. doi:10.17703/IJACT.2019.7.4.56
98. F. Rius, M. P. Philipsen, J. A. Tur, T. B. Moeslund, C. Angulo, and M. Casas, “Autoencoders for Semi-Supervised Water Level Modeling in Sewer Pipes with Sparse Labeled Data,” *Water*, vol. 14, no. 3, p. 333, Jan. 2022, doi: 10.3390/w14030333.
99. J. Jing, Z. Wang, M. Rättsch, and H. Zhang, “Mobile-U-net: An efficient convolutional neural network for fabric defect detection,” *Textile Research Journal*, vol. 92, no. 1–2, pp. 30–42, May 2020, doi: 10.1177/0040517520928604.
100. М. О. Зарецький, “Нейромережеве вимірювання рівня води в трубопроводі водовідведення за даними відеоінспекції,” *Міжнародна наукова конференція молодих учених Інформатика, Математика, Автоматика, ІМА-2023*. Суми, Україна - Астана, Казахстан, 24–28 квітня 2023, С. 45-46. Available: https://essuir.sumdu.edu.ua/bitstream-download/123456789/91550/1/Conf_IMA_2023.pdf . (accessed on: May 22, 2023).
101. J. B. Haurum and T. B. Moeslund, “Sewer-ML: A Multi-Label Sewer Defect Classification Dataset and Benchmark,” *arXiv (Cornell University)*, Mar. 2021, doi: 10.48550/arxiv.2103.10895.
102. М. О. Зарецький, “Інформаційна технологія інтелектуального аналізу даних відеоінспекції трубопроводу водовідведення,” *Міжнародна наукова конференція молодих учених Інформатика, Математика, Автоматика, ІМА-2023*. Суми, Україна - Астана, Казахстан, 24–28 квітня 2023, С. 47-48. Available: https://essuir.sumdu.edu.ua/bitstream-download/123456789/91550/1/Conf_IMA_2023.pdf . Accessed on: May 22, 2023.
103. X. Zeng, Z. Wang, H. Wang, S. Zhu, and S. Chen, “Progress in Drainage Pipeline Condition Assessment and Deterioration Prediction Models,” *Sustainability*, vol. 15, no. 4, p. 3849, Feb. 2023, doi: 10.3390/su15043849.

104. O. Alghushairy, R. Alsini, T. Soule, and X. Ma, “A Review of Local Outlier Factor Algorithms for Outlier Detection in Big Data Streams,” *Big Data and Cognitive Computing*, vol. 5, no. 1, p. 1, Dec. 2020, doi: <https://doi.org/10.3390/bdcc5010001>.

105. Z. Zhao, M. Jiang, S. Guo, Z.-Z. Wang, F. Chao, and Kay Chen Tan, “Improving Deep Learning based Optical Character Recognition via Neural Architecture Search,” *In 2020 IEEE Congress on Evolutionary Computation (CEC)* Jul. 2020, doi: [10.1109/cec48606.2020.9185798](https://doi.org/10.1109/cec48606.2020.9185798).

106. J. Park, Y.-S. Lee, E. Suh, H. Jeon, K.-J. Yoon, and K. H. Kim, “Improvement of Optical Flow Estimation by Using the Hampel Filter for Low-End Embedded Systems,” *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 7233–7239, Oct. 2021, doi: [10.1109/lra.2021.3095927](https://doi.org/10.1109/lra.2021.3095927).

107. M. Kutylowska and D. R. Kowalski, “Application of regression methods for classification of sewers’ damages,” *Applied Water Science*, vol. 11, no. 9, Sep. 2021, doi: [10.1007/s13201-021-01488-0](https://doi.org/10.1007/s13201-021-01488-0).

108. F. M. Celebi, E. A. McDaniel, T. Reite, “Creating reproducible workflows for complex computational pipelines,” *Arcadia Science*, doi:[10.57844/ARCADIA-CC5J-A519](https://doi.org/10.57844/ARCADIA-CC5J-A519)

109. D. A. Reis, B. B. Piedade, F. F. Correia, J. M. Dias, and A. Aguiar, “Developing Docker and Docker-Compose Specifications: A Developers’ Survey,” *IEEE Access*, vol. 10, pp. 2318–2329, Jan. 2022, doi: [10.1109/access.2021.3137671](https://doi.org/10.1109/access.2021.3137671).

110. K. Dergachov, L. M. Krasnov, V. Bilozerskyi, and A. Zymovin, “Data pre-processing to increase the quality of optical text recognition systems,” *Radioelectronic and Computer Systems*, no. 4, pp. 183–198, Nov. 2021, doi: [10.32620/reks.2021.4.15](https://doi.org/10.32620/reks.2021.4.15).

111. A. Radford *et al.*, “Learning Transferable Visual Models From Natural Language Supervision,” *Proceedings of the 38th International Conference on Machine Learning*, PMLR pp. 8748-8763, 2021. Available: <http://proceedings.mlr.press/v139/radford21a/radford21a.pdf>. Accessed on: May 20, 2023

112. V. Zamfir, M. Carabas, C. Carabas, and N. Tapus, “Systems Monitoring and Big Data Analysis Using the Elasticsearch System,” *In 2019 22nd International*

Conference on Control Systems and Computer Science (CSCS) doi: 10.1109/cscs.2019.00039.

113. R. Serrano, “Aws S3 como mecanismo de recuperación ante desastres tecnológicos en pymes,” *FapUNIFESP*, May 2020, doi: <https://doi.org/10.1590/scielopreprints.325>.

114. A. Pathak and C. Kalaiarasan, “RabbitMQ Queuing Mechanism of Publish Subscribe model for better Throughput and Response,” *Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT)*. 2021. doi: 10.1109/icecct52121.2021.9616722.

115. J. Sigut, M. Castro, R. Arnay, and M. Sigut, “OpenCV Basics: A Mobile Application to Support the Teaching of Computer Vision Concepts,” *IEEE Transactions on Education*, vol. 63, no. 4, pp. 328–335, Nov. 2020, doi: 10.1109/te.2020.2993013.

116. B. Pang, E. Nijkamp, and Y. Wu, “Deep Learning With TensorFlow: A Review,” *Journal of Educational and Behavioral Statistics*, vol. 45, no. 2, pp. 227–248, Sep. 2019, doi: 10.3102/1076998619872761.

117. D. Paper, *Hands-on Scikit-Learn for machine learning applications : data science fundamentals with Python*. Berkeley, Ca: Apress, 2020, pp. 47–160.

ДОДАТКИ

ДОДАТОК А

Акти впровадження результатів дисертаційної роботи

КОМУНАЛЬНЕ ПІДПРИЄМСТВО
«МІСЬКВОДОКАНАЛ»
СУМСЬКОЇ МІСЬКОЇ РАДИ
вул. Білопільський шлях, 9 м. Суми, 40009
тел. (факс) 700-181, 700-160
e-mail: vodocanal_sumy@ukr.net



SUMY CITY COUNCIL
MUNICIPAL UTILITY
«MISKVODOKANAL»
9, Bilopilskiy way, Sumy, Ukraine, 40009
tel. 700-181, 700-160
e-mail: vodocanal_sumy@ukr.net

№ 7440/18 від 26 ЛИС 2021

ДОВІДКА

про впровадження наукових результатів дисертаційної роботи
Зарецького Миколи Олександровича, виконаної на здобуття наукового ступеня доктора
філософії зі спеціальності 122 – комп'ютерні науки

Основні положення та результати наукових досліджень аспіранта Сумського державного університету Зарецького Миколи Олександровича були впровадженні і використані на КП "Міськводоканал" СМР під час підготовки персоналу, а також під час інспекції каналізаційної системи. Зокрема, було використано інформаційне та програмне забезпечення системи класифікаційного аналізу дефектів каналізаційних мереж водовідведення.

Впровадження результатів дисертаційного дослідження дозволило забезпечити повноту перевірки функціонального стану каналізаційного трубопроводу і отримати згенерований звіт, достатній для прийняття рішень щодо плану обслуговування і ремонту.

Начальник каналізаційних мереж




(підпис)

Гевліч М.О.

М.П

ТОВАРИСТВО З ОБМЕЖЕНОЮ ВІДПОВІДАЛЬНІСТЮ
«РОЗУМНІ ТЕХНОЛОГІЇ «ТИТУЛ»»

Україна, 40009, м. Суми, вул. Чернігівська, 3
Ідентифікаційний код 41779127, п/р 26007055035045 в ПАТ КБ «Приватбанк», МФО 337546
Свідоцтва про держреєстрацію №1 632 102 0000 014798 від 05.12.2017 р.
ІПН 417791218198

Вих. № _____ від 21.09.2021 р.

АКТ ВПРОВАДЖЕННЯ

наукових результатів дисертаційної роботи

Зарецького Миколи Олександровича, виконаної на здобуття наукового ступеня
доктора філософії зі спеціальності 122 – комп'ютерні науки

Основні положення наукового дослідження аспіранта Сумського державного університету Зарецького М. О., а саме :

– метод навчання ієрархічного екстрактора ознак кадрів відео моніторингу;
– інформаційне та програмне забезпечення системи аналізу даних відео моніторингу, яка функціонує за умов ресурсних та інформаційних обмежень, було використано під час розроблення мобільної інтелектуальної системи відеомоніторингу інфраструктурних об'єктів.

Впровадження результатів дисертаційного дослідження дозволило підвищити достовірність розпізнавання об'єктів інтересу під час відео моніторингу інфраструктурних об'єктів.



С. А. Светайло

Шаблон за версією 01
Затверджено наказом ректора СумДУ
№0527-1 від 14.11.2017 р.

ЗАТВЕРДЖУЮ

Перший проректор Сумського
державного університету

С.В. Леонів
(підпис) (прізвище, ініціали)

“ 5 ” лютого 2021 р.




АКТ

впровадження результатів досліджень
дисертаційної роботи Зарецького Миколи Олександровича, виконаної на здобуття ступеня
доктора філософії зі спеціальності 122 – комп'ютерні науки, у навчальний процес
Сумського державного університету

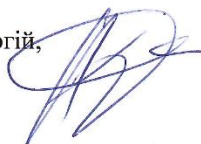
Комісія у складі: голова – декан факультету електроніки та інформаційних технологій, к.ф.-м.н., доцент О. О. Дрозденко та членів комісії: заступника декана з виховної роботи факультету електроніки та інформаційних технологій, к.ф.-м.н., старший викладач, А. А. Дворніченко і завідувача кафедри комп'ютерних наук, д.т.н., професор, А. С. Довбиш встановила, що у навчальному процесі Сумського державного університету використано такі наукові результати, одержані при виконанні дисертаційної роботи Зарецького М.О.:

- 1) результати дисертаційної роботи впроваджено у навчальний процес кафедри комп'ютерних наук Сумського державного університету при викладанні наступних дисциплін: «Розробка сервіс-орієнтованого програмного забезпечення», «Introduction to Data Science»;
- 2) результати дисертаційної роботи використано в магістерських кваліфікаційних роботах:
 - студента гр. ІН-71 Коплик А. В. на тему «Інтелектуальна система оцінки функціонального стану стічних труб»;
 - студента гр. КБ-71 Меняк А. В. на тему «Інформаційна система завадозахищеного розпізнавання контексту спостережень в задачі інспекції труб»;
 - студента гр. ІТ.м.з-92с Москаленко В. В. на тему «Інтелектуальна система розпізнавання дефектів стічних труб».

Голова комісії, декан
факультету електроніки та інформаційних
технологій, к.ф.-м.н., доцент

 О. О. Дрозденко

Члени комісії
заступник декана з виховної роботи
факультету електроніки та інформаційних технологій,
к.ф.-м.н., ст. викладач

 А. В. Дворніченко

Завідувач кафедри
комп'ютерних наук, д.т.н., професор

 А. С. Довбиш

**Molfar.AI sp. z o.o.**

Al. Grunwaldzka 472b
80-309 Gdańsk
Poland
NIP : PL5842781443
hello@molfar.ai
molfar.ai

№ 17-RB від 21.09.2021

АКТ ВПРОВАДЖЕННЯ
наукових результатів дисертаційної роботи
Зарецького Миколи Олександровича, виконаної на здобуття наукового ступеня
доктора філософії зі спеціальності 122 – комп'ютерні науки

Основні положення наукового дослідження аспіранта Сумського державного університету Зарецького М. О., а саме :

– модель і метод навчання ієрархічного екстрактора ознак кадрів відео інспекції стічних труб;

– моделі і алгоритми синтезу вирішувальних правил для класифікаційного аналізу дефектів стічних труб,

було використано під час розроблення системи класифікаційного аналізу даних відеоінспекції стічних труб для компанії MPWiK м. Вроцлав, Польща.

Впровадження результатів дисертаційного дослідження дозволило автоматично формувати звіти про функціональний стан стічних труб за даними відеоінспекції та досягнути точності прийнятної для практичного використання.

Директор компанії

О. Лисюк

21-09-2021

COMPANY NUMBER – 0000770440**Registered Address**

Al. Grunwaldzka, 472b

Gdansk, Poland

80-309

NIP / VAT No: PL 5842781443

Contact Information

Name: Elena Lysyuk

Phone: +46 704 890 560

Email: hello@molfar.ai

Website: molfar.ai

ТОВАРИСТВО З ОБМЕЖЕНОЮ ВІДПОВІДАЛЬНІСТЮ
«НОРД ТРЕК»

Україна, 41100, Сумська область, м. Шостка, вул. Весняна, буд. 36
Ідентифікаційний код 42579826, п/р UA833282090000026006000015854 в АБ «Південний», МФО 328209
Свідоцтво про держреєстрацію № 16331020000001605 від 29.10.2018 р.
ІПН № 425798218223

Вих.№61/1 від 05.05.2023 р.

АКТ ВПРОВАДЖЕННЯ
наукових результатів дисертаційної роботи
Зарецького Миколи Олександровича, виконаної на здобуття наукового
ступеня доктора філософії зі спеціальності 122 – комп'ютерні науки

Основні положення наукового дослідження аспіранта Сумського державного університету, а саме:

- модель екстракції ознакового опису даних відеоінспекції інфраструктури водовідведення з урахування просторово-часового контексту;
- інформаційне та програмне забезпечення модуля машинного зору для дистанційно-керованої мобільної платформи.

Впровадження результатів дисертаційного дослідження дозволило підвищити достовірність розпізнавання функціонального стану інфраструктури водовідведення в умовах впливу візуальних перешкод.

Директор



І. В. Кашук

ДОДАТОК Б

Список публікацій за темою дисертації

Наукові праці, в яких опубліковані основні наукові результати дисертації

Статті у фахових наукових виданнях із переліку МОН України

1. В. В. Москаленко, А. С. Москаленко, М. О. Зарецький, та А. Г. Коробов, “Модель та алгоритм навчання системи автономної навігації для малогабаритного безпілотного апарату ,” *Біоніка інтелекту*, № 91, С. 16–24, 2018.

(Особистий внесок: здобувач проаналізував існуючі алгоритми візуальної навігації).

2. В. В. Москаленко, А. С. Москаленко та М. О. Зарецький, “Модель і алгоритм навчання детектора шкідливого трафіку на основі модифікації зростаючого нейронного газу,” *Радіоелектронні І Комп’ютерні Системи*, № 3, С. 11–19, 2018, doi: 10.32620/reks.2018.3.02.

(Особистий внесок: здобувач проаналізував існуючі алгоритми синтезу інформаційно-екстремальних вирішувальних правил).

3. В.В. Москаленко, А. А. Москаленко, А.Г. Коробов, М. О. Зарецький, та В.А. Семашко, “Модель та алгоритм навчання системи детектування малорозмірних об’єктів для малогабаритних безпілотних літальних апаратів,” *Радіоелектронні І Комп’ютерні Системи*, № 4, С. 41–52, 2018, doi: 10.32620/reks.2018.4.04.

(Особистий внесок: здобувач проаналізував існуючі алгоритми екстракції ознакового опису спостережень для детектування малорозмірних об’єктів інтересу на зображенні).

4. В. В. Москаленко, М. О. Зарецький, та А. А. Москаленко, “Модель та метод навчання системи класифікації рухомих об’єктів для малогабаритного безпілотного апарату,” *Радіоелектронні І Комп’ютерні Системи*, №. 2, С. 108–117, 2019, doi: 10.32620/reks.2019.2.10.

(Особистий внесок: здобувач проаналізував існуючі алгоритми машинного навчання для розпізнавання об’єктів інтересу в процесі руху).

Статті у наукових періодичних виданнях інших держав та виданнях, які індексуються наукометричними базами даних Scopus та Web of Science

5. V. V. Moskalenko, A. S. Moskalenko, A. G. Korobov, and M. O. Zaretsky, “A Model and Training Algorithm of Small-Sized Object Detection System for a Compact Aerial Drone,” *Radio Electronics, Computer Science, Control*, vol. 0, no. 1, Apr. 2019, doi: <https://doi.org/10.15588/1607-3274-2019-1-11>. (Web of Science)

(Особистий внесок: здобувач проаналізував існуючі алгоритми детектування об’єктів інтересу на зображенні за умов ресурсних та інформаційних обмежень).

6. В. В. Москаленко, М. О. Зарецький, А. А. Москаленко, А. А. Кудрявцев, and В. Ю. Семашко, “Багатошарова модель та метод навчання для детектування шкідливого трафіку на основі ансамблю дерев рішень,” *Радіоелектронні І Комп’ютерні Системи*, № 2, С. 92–101, Apr. 2020, doi: 10.32620/reks.2020.2.08. (Scopus)

(Особистий внесок: здобувач розробив алгоритм оцінювання ефективності інформаційно-екстремальних вирішувальних правил).

7. V. V. Moskalenko, M. O. Zaretskyi, A. S. Moskalenko, and V. Lysyuk, “Sewer Pipe Defects Classification Based on Deep Convolutional Network with Information-Extreme Error-Correction Decision Rules,” in *Springer eBooks*, 2020, pp. 253–263. doi: 10.1007/978-3-030-61656-4_16. (Scopus)

(Особистий внесок: здобувач розробив модель та алгоритм навчання для розпізнавання дефектів труб водовідведення, коли камера орієнтована вперед вздовж труби, а також виконав експериментальні дослідження).

8. В. В. Москаленко, М. О. Зарецький, Я. Ю. Ковальський, та С. С. Мартиненко, “Модель і метод навчання класифікатора контекстів спостереження на зображеннях відеоінспекції стічних труб,” *Радіоелектронні І Комп’ютерні Системи*, № 3, С. 59–66, 2020, doi: 10.32620/reks.2020.3.06. (Scopus)

(Особистий внесок: здобувач сформував вхідний математичний опис та розробив глибоку модель з інформаційно-екстремальними вирішувальними

правилами для розпізнавання контексту спостереження на кадри відеоінспекції, а також виконав експериментальні дослідження).

9. В. В. Москаленко, М. О. Зарецький, А. С. Москаленко, А. Г. Коробов, та Я. Ю. Ковальський, “Багатоетапний метод глибинного навчання з попереднім самонавчанням для класифікаційного аналізу дефектів стічних труб,” *Радіоелектронні І Комп’ютерні Системи*, № 4, С. 71–81, 2021, doi: 10.32620/reks.2021.4.06. (Scopus)

(Особистий внесок: здобувач розробив багатоетапний метод навчання глибокої моделі з ініціалізацією вагових коефіцієнтів в процесі контрастного самонавчання для розпізнавання дефектів труб водовідведення, коли камера орієнтована вбік до стінки труби).

10. В. В. Москаленко, М. О. Зарецький, А. Г. Коробов, Я. Ю. Ковальський, А. Ф. Шаєхов, В. А. Семашко, та ін., “Модель та метод навчання для класифікаційного аналізу рівня води в стічних трубах за даними відео інспекції,” *Радіоелектронні І Комп’ютерні Системи*, № 2, С. 4–15, 2021, doi: 10.32620/reks.2021.2.01. (Scopus)

(Особистий внесок: здобувач сформував вхідний математичний опис, розробив глибоку модель для розпізнавання рівня води в трубі на зображеннях відео інспекції з урахуванням міжкадрових залежностей, а також виконав експериментальні дослідження).

11. V. V. Moskalenko, M. O. Zaretsky, A. S. Moskalenko, A. O. Panych, and V. V. Lysyuk, “A Model and Training Method for Context Classification in Cctv Sewer Inspection Video Frames,” *Radio Electronics, Computer Science, Control*, no. 3, pp. 97–108, Oct. 2021, doi: <https://doi.org/10.15588/1607-3274-2021-3-9>. (Web of Science)

(Особистий внесок: здобувач сформував вхідний математичний опис, розробив глибоку модель для розпізнавання контексту спостереження на зображеннях відео інспекції з урахуванням міжкадрових залежностей, а також виконав експериментальні дослідження).

12. V. V. Moskalenko, A. S. Moskalenko, A. G. Korobov, and M. O. Zaretsky, “Image Classifier Resilient to Adversarial Attacks, Fault Injections and Concept Drift –

Model Architecture and Training Algorithm,” *Radio Electronics, Computer Science, Control*, no. 3, p. 86, Oct. 2022, doi: 10.15588/1607-3274-2022-3-9. (Web of Science)

(Особистий внесок: програмно реалізував згорткову нейронну мережу з інформаційним пляшковим горлом та гіперсферичними контейнерами класів розпізнавання).

Наукові праці, які засвідчують апробацію матеріалів дисертації

13. В. В. Москаленко, М. О. Зарецький, Д. О. Лєвокутник, “Модель і метод навчання детектора малорозмірних об’єктів за умов ресурсних та інформаційних обмежень,” *XIV International scientific conference «Intellectual systems of decision-making and problems of computational intelligence»*. с. Залізний Порт, 21-25 трав. 2019 С. 116-118

(Особистий внесок: здобувач проаналізував існуючі методи машинного навчання глибоких моделей детектування об’єктів).

14. V. V. Moskalenko, A. S. Moskalenko, and M. O. Zaretskyi, “Model and training method of moving object classification system for a compact drone,” *14th International Conference on Computer Sciences and Information Technologies (CSIT)*, 2019, pp. 56–61. doi: 10.1109/stc-csit.2019.8929795. (Scopus & Web of Science)

(Особистий внесок: здобувач проаналізував існуючі алгоритми детектування об’єктів інтересу на місцевості в процесі руху).

15. M. Zaretskyi V. Moskalenko, A. Moskalenko, and A. Shaiekhov, “Multi-Layer Model and Training Method for Information-Extreme Malware Traffic Detector,” *The Third International Workshop on Computer Modeling and Intelligent Systems*, April 27-May 1, 2020, pp. 288-299 Available at: <https://ceur-ws.org/Vol-2608/paper22.pdf>. (accessed on May 1 2021). (Scopus)

(Особистий внесок: здобувач запропонував метод обчислення еталонного вектора кожного класу розпізнавання в двійковому просторі для синтезу інформаційно-екстремальних вирішувальних правил).

16. V. V. Moskalenko, A.S. Moskalenko, M.O. Zaretskyi, and V. Lysyuk, “Deep Feature Extractor with Information-Extreme Decision Rules for Visual Classification of Sewer Pipe Defects and its Training Method,” *Data Stream Mining & Processing*, pp. 191–195, 2022 doi: 10.1109/dsmp47368.2020.9204083. (Scopus & Web of Science)

(Особистий внесок: здобувач сформував вхідний математичний опис та розробив глибоку модель з інформаційно-екстремальними вирішувальними правилами для розпізнавання дефектів у випадку, коли камера орієнтована вперед вздовж труби).

17. М. О. Зарецький, “Нейромережеве вимірювання рівня води в трубопроводі водовідведення за даними відеоінспекції,” *Міжнародна наукова конференція молодих учених Інформатика, Математика, Автоматика, ІМА-2023*. Суми, Україна - Астана, Казахстан, 24–28 квітня 2023, С. 45-46. Available: https://essuir.sumdu.edu.ua/bitstream-download/123456789/91550/1/Conf_IMA_2023.pdf . (accessed on: May 22, 2023).

(Особистий внесок: здобувач проаналізував існуючі методи машинного навчання глибоких моделей детектування об’єктів).

Наукові праці, які додатково відображають наукові результати дисертації

18. В.В. Москаленко, А.С. Москаленко, М.О. Зарецький, В.В. Нагорний, Свідоцтво про реєстрацію авторського права на твір № 89337, комп’ютерна програма «Інформаційно-екстремальний класифікатор багатовимірних спостережень з метаевристичною оптимізацією параметрів функціонування» заяв. 2019-04-05; опубл. 2019-06-05.

(Особистий внесок: програмна реалізація фітнес-функції мета-евристичного алгоритму, в якій обчислюється інформаційний критерій ефективності моделі аналізу даних на тестовій вибірці)

19. В.В. Москаленко, А.С. Москаленко, М.О. Зарецький, А.О. Панич, Свідоцтво про реєстрацію авторського права на твір № 106341, комп’ютерна

програма «Програма для розпізнавання та видалення текстових анотацій з відеозображень» заяв. 2021-07-05; опубл. 2021-07-16.

(Особистий внесок: програмна реалізація функції заповнення виділеної ділянки зображення на основі інформації про оточуючі ділянки зображення)

20. В.В. Москаленко, А.С. Москаленко, М.О. Зарецький, А.О. Панич, А.Г. Коробов, Свідоцтво про реєстрацію авторського права на твір № 107201, комп'ютерна програма «Реалізація алгоритму інформаційно-екстремального навчання глибокої моделі аналізу даних з використанням принципів сіамських мереж та завадозахищеного кодування» дата реєстр. 11.08.2021.

(Особистий внесок: здобувач виконав програмну реалізацію алгоритму навчання глибокої згорткової моделі для побудови інформаційно-екстремального класифікатора зображень відеоінспекції стічних труб).

21. В.В. Москаленко, А.С. Москаленко, М.О. Зарецький, А.Г. Коробов, Свідоцтво про реєстрацію авторського права на твір № 113206 «Програма машинного навчання класифікатора зображень з підвищеною робастністю до шуму і змагальних атак» заяв. 2022-06-06; опубл. 2022-07-29.

(Особистий внесок: програмна реалізація алгоритму аугментації даних з використанням шумових шаблонів)