Edited by
Volodymyr Lukin

PATTERN RECOGNITION IN SURVEILLANCE SYSTEMS AND DIAGNOSTICS

Collective monograph

Scientific Route®

Tallinn

Published in June 2026
by Scientific Route OÜ®
Parda tn 4, Kontor 526, Tallinn, Harju maakond Estonia, 10151
www.route.ee

Pattern recognition in surveillance systems and diagnostics
Volodymyr Lukin (Editor)

DOI: 10.21303/978-9908-845-05-0
ISBN 978-9908-845-05-0 (eBook)
ISBN 978-9908-845-06-7 (ePub)

This publication has been peer reviewed.

ISBN 978-9908-845-05-0 (eBook)
ISBN 978-9908-845-06-7 (ePub)

AUTHORS

Chapter 1

Volodymyr Lukin
Doctor of Technical Sciences, Professor,
Head of Department
Department of Information and Communication Technologies
National Aerospace University "Kharkiv Aviation Institute"
ORCID: https://orcid.org/0000-0002-1443-9685

Sergii Kryvenko
Doctor of Technical Sciences, Senior Researcher
Department of Information and Communication Technologies
National Aerospace University "Kharkiv Aviation Institute"
ORCID: https://orcid.org/0000-0001-6027-5442

Fangfang Li
Doctor of Philosophy (PhD), Lecturer
Key Laboratory of Opto-Electronic Information Science and Technology of Jiangxi Province
Nanchang Hangkong University
ORCID: https://orcid.org/0000-0002-8212-9144

Sergiy Abramov
Candidate of Technical Sciences,
Associate Professor, Chief Engineer
Department of Functional Materials and Electronics
Center for Physical Sciences and Technology
ORCID: https://orcid.org/0000-0002-8295-9439

Viktoriia Abramova
Candidate of Technical Sciences, Researcher
Department of Optoelectronics
Center for Physical Sciences and Technology
ORCID: https://orcid.org/0000-0001-5802-5858

Bohdan Kovalenko
Doctor of Philosophy (PhD), Senior Lecturer
Department of Information and Communication Technologies
National Aerospace University "Kharkiv Aviation Institute"
ORCID: https://orcid.org/0000-0002-9360-0691

Igor Dohtiev
PhD Student
Department of Information and Communication Technologies
National Aerospace University "Kharkiv Aviation Institute"
ORCID: https://orcid.org/0009-0000-8510-958X

Oleksandr Arkhipov
PhD Student
Department of Information and Communication Technologies
National Aerospace University "Kharkiv Aviation Institute"
ORCID: https://orcid.org/0009-0009-6943-0870

Nenad Stojanović
Doctor of Philosophy (PhD), Teaching Assistant
Department of Telecommunications and Informatics
Military Academy
University of Defence
ORCID: https://orcid.org/0000-0001-9328-5348

Boban Bondžulić
Doctor of Philosophy (PhD), Associate Professor
Department of Telecommunications and Informatics
Military Academy
University of Defence
ORCID: https://orcid.org/0000-0002-8850-9842

Chapter 2

Pavlo Mykhalichenko
Doctor of Technical Sciences, Associate Professor
Department of Automatics and Electrical Equipment
Kherson Educational-Scientific Institute
of Admiral Makarov National University
of Shipbuilding
ORCID: https://orcid.org/0000-0002-2680-140X

Tetiana Cherniavska
Doctor of Economic Sciences, Professor
Department of Economics and Technical Sciences
State University of Applied Sciences in Konin
ORCID: https://orcid.org/0000-0002-4729-2157

Bohdan Cherniavskyi
Doctor of Technical Sciences, Adjunct
Department of Economics and Technical Sciences
State University of Applied Sciences in Konin
ORCID: https://orcid.org/0000-0001-9174-6139

Viktor Nadtochii
Candidate of Technical Sciences,
Associate Professor
Department of Automatics and Electrical Equipment
Kherson Educational-Scientific Institute
of Admiral Makarov National University
of Shipbuilding
ORCID: https://orcid.org/0000-0003-3869-3546

Anatolii Nadtochyi
Candidate of Technical Sciences,
Associate Professor,
Acting Head of Department
Department of Automatics and Electrical Equipment
Kherson Educational-Scientific Institute
of Admiral Makarov National University
of Shipbuilding
ORCID: https://orcid.org/0000-0003-1905-0895

Maiia Korkh
Candidate of Pedagogical Sciences,
Associate Professor
Department of Engineering Mechanics
and Ship Repair
National University "Odessa Maritime Academy"
ORCID: https://orcid.org/0000-0002-9401-5240

Chapter 3

Tetiana Cherniavska
Doctor of Economic Sciences, Professor
Department of Economics and Technical Sciences
State University of Applied Sciences in Konin
ORCID: https://orcid.org/0000-0002-4729-2157

Bohdan Cherniavskyi
Doctor of Technical Sciences, Adjunct
Department of Economics and Technical Sciences
State University of Applied Sciences in Konin
ORCID: https://orcid.org/0000-0001-9174-6139

Oksana Zghurska
Doctor of Economic Sciences, Associate Professor,
Head of Department
Department of Economics and Entrepreneurship
State University of Information and Communication
Technologies
ORCID: https://orcid.org/0000-0003-3878-3007

Serhii Kasian
Candidate of Economic Sciences,
Associate Professor
Department of Marketing
Dnipro University of Technology
ORCID: https://orcid.org/0000-0002-7103-4457

Kateryna Nakonechna
Candidate of Economic Sciences,
Associate Professor
Department of Economic Cybernetics
National University of Life and Environmental
Sciences of Ukraine
ORCID: https://orcid.org/0000-0002-1537-7201

Yaroslava Mudra
Candidate of Economic Sciences,
Associate Professor
Department of International Economic Relations
Khmelnytskyi National University
ORCID: https://orcid.org/0000-0001-7343-6735

Chapter 4

Tetiana Cherniavska
Doctor of Economic Sciences, Professor
Department of Economics and Technical Sciences
State University of Applied Sciences in Konin
ORCID: https://orcid.org/0000-0002-4729-2157

Bohdan Cherniavskyi
Doctor of Technical Sciences, Adjunct
Department of Economics and Technical Sciences
State University of Applied Sciences in Konin
ORCID: https://orcid.org/0000-0001-9174-6139

Alla Rusnak
Doctor of Economic Sciences, Professor,
Acting Head of Department
Department of Economics
Kherson Educational-Scientific Institute
of Admiral Makarov National University
of Shipbuilding
ORCID: https://orcid.org/0000-0002-3198-2866

Iryna Nadtochii
Doctor of Economic Sciences, Professor,
Associate Director Education and Research
Kherson Educational-Scientific Institute
of Admiral Makarov National University
of Shipbuilding
ORCID: https://orcid.org/0000-0003-0693-8000

Artur Harahulia
Doctor of Philosophy (PhD)
Department of Economics
Kherson Educational-Scientific Institute
of Admiral Makarov National University
of Shipbuilding
ORCID: https://orcid.org/0000-0001-5829-2112

Oleksii Bobrovskyi
Doctor of Philosophy (PhD)
Department of Economics
Kherson Educational-Scientific Institute
of Admiral Makarov National University
of Shipbuilding
ORCID: https://orcid.org/0000-0001-7395-7477

Chapter 5

Dmytro Krytskyi
Candidate of Technical Sciences,
Associate Professor, Dean
Department of Information Technologies of Design
National Aerospace University "Kharkiv Aviation
Institute"
ORCID: https://orcid.org/0000-0003-4919-0194

Elvira Kaidan
Department of Information Technologies of Design
National Aerospace University "Kharkiv Aviation
Institute"
ORCID: https://orcid.org/0009-0007-1437-6884

Artem Chekhovsky
PhD student
Department of Information Technologies of Design
National Aerospace University "Kharkiv Aviation
Institute"
ORCID: https://orcid.org/0009-0005-9704-0504

Chapter 6

Yevgeniy Kalinichenko
Candidate of Technical Sciences,
Associate Professor, Head of Department
Department of Navigation and Ship Control
Odesa National Maritime University
ORCID: https://orcid.org/0000-0003-2898-7313

Oleksandr Koliesnik
Senior Lecturer
Department of Navigation and Ship Control
Odesa National Maritime University
ORCID: https://orcid.org/0009-0003-3713-2015

Oleg Safyan
Senior Lecturer
Department of Navigation and Ship Control
Odesa National Maritime University
ORCID: https://orcid.org/0000-0001-8866-7456

Chapter 7

Bohdan Cherniavskyi
Doctor of Technical Sciences, Adjunct
Department of Economics and Technical Sciences
State University of Applied Sciences in Konin
ORCID: https://orcid.org/0000-0001-9174-6139

Oksana Drozd
Candidate of Technical Sciences,
Associate Professor
Kherson Educational-Scientific Institute
of Admiral Makarov National University
of Shipbuilding
ORCID: https://orcid.org/0000-0002-0135-8659

Anatolii Nadtochyi
Candidate of Technical Sciences,
Associate Professor, Acting Head
of Department
Department of Automatics and Electrical
Equipment
Kherson Educational-Scientific Institute
of Admiral Makarov National University
of Shipbuilding
ORCID: https://orcid.org/0000-0003-1905-0895

Viktor Nadtochii
Candidate of Technical Sciences,
Associate Professor
Department of Automatics and Electrical
Equipment
Kherson Educational-Scientific Institute
of Admiral Makarov National University
of Shipbuilding
ORCID: https://orcid.org/0000-0003-3869-3546

Maksym Matviienko
Candidate of Technical Sciences,
Associate Professor
Department of Welding
Kherson Educational-Scientific Institute
of Admiral Makarov National University
of Shipbuilding
ORCID: https://orcid.org/0000-0002-1020-0415

Ivan Kalinichenko
Candidate of Technical Sciences,
Associate Professor
Department of Thermal Engineering
Kherson Educational-Scientific Institute
of Admiral Makarov National University
of Shipbuilding
ORCID: https://orcid.org/0000-0001-6765-6168

ABSTRACT

The collective monograph "Pattern recognition in surveillance systems and diagnostics" deals with recent advances in several related areas that are rapidly developing and need efficient state-of-the-art solutions. Although particular tasks solved by chapters' authors are quite diverse, they are based on modern approaches of digital signal/image processing, artificial intelligence, training, and further verification, adaptation, and optimization.

Quality of lossy compressed images and ways of its providing

This chapter considers the problem of providing a desired quality of images compressed in a lossy manner due to the necessity to radically decrease the amount of registered, transferred, or disseminated data without significant reduction of efficiency in solving object recognition or sensed terrain classification tasks. Visually lossless compression is treated as a possible solution where a quite large compression ratio and appropriate quality of compressed images can be ensured simultaneously. Such an approach can also be useful for medical diagnostic systems.

Solving pattern recognition problems in shipping monitoring systems based on CNN-models

The chapter proposes a fundamentally new convolutional neural network (CNN) architecture for vision-based monitoring and classification of inland waterway vessels in river-port environments. A comprehensive review of the fundamental principles of CNN construction is provided. More than 12,000 annotated images of inland waterway vessels have been collected in a dataset employed in experimental evaluation, according to which the proposed CNN model achieved the F1-score of 91.9%. It has outperformed classical methods by 8–12% on average.

Reengineering of management processes for the restoration of transport and logistics infrastructure through image recognition and BIM-oriented remediation

This chapter focuses on aspects of the reengineering of infrastructure facilities. The authors expand the range of application of computer vision, analysis of images and UAV imagery, satellites, and neural-network recognition, which is integrated with the BIM model of the facility, forming a "digital twin". This makes it possible to perform more accurate planning of remediation/reconstruction/restoration works (4D/5D), choose the optimal scenario of necessary measures, etc. The study also makes it possible to assess the effect of the implemented complex of works.

From pattern recognition to remediation management in a closed digital loop for post-war logistics

The authors study the problem of transition from digital diagnostics to executable management of remediation processes for territories influenced by emergency events, including military actions. They show that the effectiveness of recovery measures is strictly connected with the quality of logistical coordination and the speed of managerial decision-making. The necessity of transitioning from the fragmented use of digital tools to an integrated management architecture is stated. The architecture of the Closed Digital Diagnostic Loop for Remediation Logistics is determined where the Pattern Recognition block acts as a source of diagnostic events, forming the basis for geospatial integration.

Detection of small and camouflaged objects

This chapter deals with an important practical task of detecting small and camouflaged objects in images acquired from UAVs, drones, and other aerial carriers. Such objects are often masked or obscured by terrain texture or other objects, leading to small objects missing. The authors focus on YOLO-based detectors due to their appropriate detection abilities and high processing speed. An experiment is carried out for an annotated dataset. The obtained results show that the improved approach ensures more stable detection for complex scenes and simultaneously provides significantly reduced training time. Among the considered CNNs, the YOLOv11 model has demonstrated the best results in detecting small targets in complex background conditions.

Visual pattern recognition in navigation simulator interfaces: a method for automatic restoration reconstruction of vessel motion parameters

In this chapter, the authors tested a simple idea: whether it is possible to obtain the ship's trajectory and navigation parameters from the image of the deck simulator interface. The corresponding image recognition methods have been designed and tested. The simulator display is treated as a reflection of the internal state of the model, where the display indicators show the state of the boat. The analysis allows representing the change in the vessel parameters over time. This also allows investigating ship dynamics and supporting experimental studies.

BPMN as a tool for adaptive support of pattern recognition and digital diagnostic results in remediation and post-crisis recovery

The chapter develops an approach to the use of BPMN as a means for support of pattern recognition and digital diagnostic results with application to remediation and post-crisis recovery. The research is important for practice since it considers the possibility of the construction of managed digital response workflows where the results of automated detection are converted into a set of coordinated actions

having the goal to eliminate the crisis consequences. The thorough verification of this approach is possible using a combination of open benchmark datasets employed for the recognition of damage and changes.

Keywords

Image and video data, visual quality, metrics, lossy compression, computational efficiency, convolutional neural networks, pattern recognition, vessel classification, shipping monitoring systems, machine learning, digital twins, transport, logistics, post-war remediation, emergency situations, closed digital diagnostic-managerial loop, geographic information systems, small and camouflaged object detection, YOLO, dataset annotation, automated data collection, ship dynamics, crisis impact, practical verification.

CIRCLE OF READERS AND SCOPE OF APPLICATION

The collective monograph "Pattern recognition in surveillance systems and diagnostics" deals with a wide scope of topics and is addressed to a broad area of researchers and practitioners in digital signal/image processing, remote sensing, pattern recognition, transport management, decision-making, telecommunications, machine learning, artificial intelligence, automation and robotic systems, etc. It can also be helpful for engineers, programmers, and regional authorities.

The research findings presented in this monograph include the following:

– *approaches to adaptive lossy compression by different techniques* that can be applied to images of different origins to produce a desired visual quality or invisibility of introduced distortions with appropriate accuracy to ensure significant reduction of data volume without loss of data value in the sense of classification accuracy, object recognition, and preservation of diagnostically important features; the application areas are remote sensing, medicine, and telecommunications;

– *modern methods of shipping monitoring and control in complex conditions* that are based on convolutional neural networks trained to diminish the negative influence of high variability of vessel shapes, viewing angles, illumination conditions, and significant intra-class differences; this has led to significant improvement of recognition accuracy and diminishing the shipping risks; the application area is shipping monitoring;

– *original approaches to restoration of transport and logistics infrastructure* oriented on eliminating the consequences of military conflicts, disasters, and man-made accidents; the approaches are data-driven in the sense that they jointly exploit computer vision, analysis of images acquired from UAVs and satellites, and neural-network recognition methods with forming a "digital twin"; the main application area is transport and infrastructure recovery;

– *means for planning remediation actions* based on digital diagnos tics and intended for management of remediation processes for territories suffering from emergency events such as military actions or infrastructure damage; the main application area is reconstruction of Ukraine in the post-war period or similar situations in other countries;

– *methods and tools for detection and localization of small and camouflaged objects* based on CNN training for appropriate datasets with marked objects of interest in conditions close to real life; the advantage of the proposed approach is the ability to perform detection tasks quite quickly with limited power consumption; the application areas are the development of smart cities and mine detection for territories after wars;

– *methods of intellectual image processing for navigation control* based on acquiring image information from the simulator screen and retrieving navigation parameters from such image data; the advantage of these methods is their simplicity and appropriate accuracy simultaneously that allows investigating and controlling ship dynamics; the application area is shipping control and management in ports and other regions of intensive traffic;

– *tools for adaptive support of pattern recognition and digital diagnostics* that presume processing of uncertainty associated with the probabilistic nature of AI outputs using the confidence thresholds, escalation mechanisms, repeated data collection, and feedback; the tools attempt to remove the gap between the digital detection of a problem and the actual implementation of recovery measures and can be applied for post-war recovery.

Therefore, the monograph can serve as a source of information about recent trends in pattern recognition and diagnostics and a guide for the development of systems and tools for solving specific practical tasks.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Modern trends in pattern recognition and applications

Pattern recognition can be considered from different perspectives: historical, methodological, technological, and application-oriented. Literature sources give different date sequences of milestones in the development of pattern recognition theory. Some of them start from philosophical and psychological sources of the beginning of the 20th century. However, all the sources mark the period of 1930–1950 as forming the mathematical basis and raising the key ideas in cybernetics, facilitating the interest in pattern recognition in its simplest forms. It was understood that animals and humans could solve a lot of tasks due to processing a particular type of information (signals in a wide sense) or several types of information jointly. People started to ask themselves whether it was possible to realize some actions similarly to animals and humans by appropriate processing of analog and, later, digital signals. Due to joint efforts and the fast development of several areas of science, pattern recognition started to become a particular direction that manifested itself by the appearance of famous books of classic texts by Duda and Hart in 1973, Fukunaga in 1972, and some others (many of them were reissued later with the incorporation of new knowledge and achievements).

Any pattern recognition system has, first, to register and/or collect data from available sources and then carry out a certain feature extraction and, probably, their representation or preparation; after this, similarity detection is to be performed using a designed classifier. In the process of such a design, performance evaluation is a typical step. At each of these stages, there are numerous factors affecting how well the final pattern recognition goal is reached. Consider an urgent task of UAV or drone detection, localization, and classification. At least, there are four sources that can be potentially employed: vision, sound (acoustic signal), reflections of radar signals, and infrared (thermal) irradiation. Meanwhile, not all sources of information can be available simultaneously, infrared data are much more informative during nighttime, while visual observation is able to give more information during daytime, although in good weather conditions.

However, even if potential sources of information for pattern recognition are clear or can be established, the next stage task is to retrieve useful features and

 DOI: 10.21303/978-9908-845-05-0.introduction

represent them properly. For the application mentioned above, this can be a drone image in visual or infrared bands (if resolution is appropriate), the object shape and color, the spectrum of acoustic signal, a radar portrait, or information on the object speed retrieved from radar reflection, etc. Then, the general feature extraction tasks divide into subtasks of finding the best features, selecting their number, providing real-time processing if necessary, combining the features of different origins, etc.

Similar problems often arise in other applications of pattern recognition. One good example is content-based image retrieval (CBIR) – the research topic that was popular about 30 years ago, the achievements of which were later realized in many modern browsers. At first sight, it is clear that shape, color, and texture features should be exploited. However, first, later it has occurred that image statistics can be useful, and, second, the tasks of finding the best feature representation, choosing their number, and the best aggregation strategy appear.

One more aspect is the rapid development of some scientific areas that, for a certain period, become a trend, seeming to be able to solve "earlier unsolvable" tasks. In the 90s of the previous centuries, it seemed that wavelets would throw other orthogonal transforms into the past and neural networks would replace all earlier designed classifiers. However, after one or two decades of accelerated development, the studies often slow down and attention is switched to other directions (methods, tools). At the moment, there are numerous classifiers, including not only neural network-based ones but also support vector machine, decision tree, and statistical classifiers.

Depending upon a situation at hand, different classifiers might occur as the best. This depends on a task to be solved, available a priori information, criteria of optimality or metrics used for performance evaluation, learning strategy, etc. A tendency of recent years is to employ convolutional neural networks (CNNs), the main advantages of which are the ability to learn better than conventional NNs due to better extraction of features. Numerous publications have appeared showing the applicability of approaches based on deep learning and artificial intelligence to solving various tasks in absolutely different areas (image processing, remote sensing, economics, agriculture, etc.). Some particular cases are studied in this monograph.

Three critical problems in deep learning-based data processing for pattern recognition and diagnostics are the following. First, one needs to have a database (dataset) or databases for learning with previously marked objects (classes). These datasets have to be large enough (e.g., contain hundreds or thousands of images with objects having different sizes, orientations, viewing conditions, and so on). This explains why numerous databases have appeared recently and continue to be designed (e.g., for

infrared and underwater images). Second, there are many types and modifications of CNNs, and they continue to appear. Then, it is a problem to choose the best CNN for an application at hand or to retrain the already used CNN if conditions of its use have changed. Third, even for the best CNN trained well, it is always possible that the pattern recognition (classification) result is wrong. A question then is, what are the consequences of such wrong recognition, especially taking into account that it is becoming more and more popular to deliver decision making to artificial intelligence?

Summarizing the aforesaid, it is possible to state the following:

1) pattern recognition continues to develop, encapsulating new areas of its application;

2) it is possible to expect rapid development of machine learning and AI approaches for at least the next decade;

3) pattern recognition and diagnostic approaches become more motivated intuitively (based on assumptions and achievements of studies of human brain activity) rather than mathematically; due to this, learning is put into basis, and, thus, the creation of the corresponding databases becomes necessary as well as significant computations;

4) there is a caution that any trained CNN is able to produce quite reliable results only for conditions for which it was properly trained, while the outcomes for unexpected situations (conditions) can be wrong with high probability, and this can cause problems.

CHAPTER 1

Quality of lossy compressed images and ways of its providing

Volodymyr Lukin
Sergii Kryvenko
Fangfang Li
Sergiy Abramov
Viktoriia Abramova
Bohdan Kovalenko
Igor Dohtiev
Oleksandr Arkhipov
Nenad Stojanović
Boban Bondžulić

Abstract

A task of lossy compression of images with providing a desired quality is considered. Image properties and metrics that are able to characterize compressed image quality including visual quality are briefly discussed. It is mentioned that a reasonable compromise between compression ratio and quality has to be provided in practice where a larger compression ratio usually leads to worse quality according to different criteria. Examples of basic dependencies are given and it is shown that the basic characteristics of compression considerably depend on image complexity that can be described by entropy or some other parameters. In addition, the compromise depends on a coder used and a parameter it employs for compression control. Analysis of basic rate-distortion curves is carried out and it is demonstrated that a desired quality according to a given metric is provided for a wide range of values of a parameter that controls compression (PCC) for a given coder. Then, the corresponding PCC has to be provided for a considered coder. Several ways to do this are considered. Drawbacks of setting a fixed PCC are discussed. A two-step approach is described and its basic properties are considered. Advantages and disadvantages of this approach are presented. An approach based on just noticeable difference is described. Its positive features and restrictions are presented. Finally, conclusions are given.

 DOI: 10.21303/978-9908-845-05-0.ch1

Keywords
Lossy compression, images, quality control, rate, distortion curve, accuracy.

1.1 Introduction

Lossy compression of images is currently widely employed in different applications including remote sensing [1], medical diagnostics [2], military [3], forestry and agriculture [4, 5], ecological monitoring [6], and so on. On one hand, images of different origin are able to provide useful information about sensed terrains, objects, and phenomena where imaging is a convenient way to have pre-requisites to obtain and retrieve such information from acquired data. On the other hand, average size of modern images and their number have an obvious tendency to increasing that makes problematic the image data transfer, storage, and dissemination. This stimulated the need in image compression where lossless compression, although being still used in some applications [7], becomes less popular [8]. The main reason is that lossy compression is able to provide a significantly larger compression ratio (CR) than for lossless compression (up to tens and hundreds [9]), that can be varied.

Meanwhile, lossy compression leads to distortions (losses, degradations) that can have different level and properties depending on many factors [9–13]. The main of them are complexity of an image subject to compression [10, 11, 13], a used compression technique and a value of a PCC [9, 13] (e.g., quality factor (QF) serves as PCC in JPEG), noise presence or absence [12], etc. A general tendency is that the introduced losses increase and a compressed image quality diminishes if CR becomes larger (although there are specific exceptions relating to the so-called strange images [14]). Then, a question arises what is an appropriate compromise between CR and image quality [8, 9, 15–17]? An answer to this question depends on a goal a compressed image is used. A compressed image can be a subject for visualization to perceive or analyze it. Then, visual quality is of prime importance and visually lossless compression can be helpful [9, 17, 18]. A compressed image can be also employed for classification or recognition [4, 19, 20]. Then, a task is to prevent significant (inappropriate) degradation of classification and object recognition accuracy due to lossy compression.

The next question is how this compromise can be attained in practice. An answer to it again depends on a coder used, chosen approach or metric to characterize compressed image quality, noise presence or absence, image complexity, and requirements to accuracy to image quality providing and computational complexity [9–13]. The size of an image to be compressed and a number of channels (components)

in this image might have an impact too. Therefore, a plethora of studies has to be carried out or the results under interest have to be retrieved from different papers to answer these questions. Below, let's present some results described in our recent papers as well as consider the outcomes and conclusions from the papers of other authors to give an insight on theoretical and practical solutions available for providing a desired quality of lossy compressed images.

1.2 Considered images and their properties

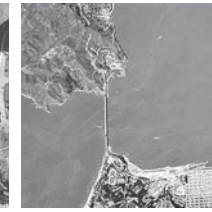
Images employed in different applications might be acquired by different types of sensors and, thus, their main properties (size, number of components, complexity, noisy or not) can vary within wide limits. Images containing millions of pixels can be met in practice of remote sensing, medical diagnostics, and customer digital photos very often. Whilst single- (e.g., medical or synthetic aperture radar) and three-channel (e.g., color) images are, probably, the most common, two-channel (dual-polarization) data and images with about ten (multispectral) and even hundreds (hyperspectral) of components are good candidates for lossy compression, too. In the latter two cases, the image size can be especially large [8, 21, 22].

Concerning image complexity, different approaches to describe it are possible. In particular, image entropy and many other parameters are able to reflect complexity of practically noise-free images [12, 23–26]. For example, **Table 1.1** presents four grayscale images of different complexity characterized by entropy E (the smaller, the simpler the image structure) and mean square error (MSE) values for distortions introduced by a better portable graphics (BPG) coder [27] for which Q (an integer-valued parameter from 0 to 51) serves as PCC where a larger Q corresponds to greater distortions and a larger CR. For smaller entropy, a smaller MSE is observed where for both values of Q the introduced distortions are invisible (this happens if distortions are similar to additive white Gaussian noise (AWGN) [11] and its variance is less than 20 for 8-bit grayscale images).

For noisy images, i.e. images for which noise is visible, the situation with characterizing their complexity is more complicated [23]. Many parameters applicable for characterization of complexity for noise-free images are sensitive to noise. Due to this, only a limited number of parameters able to do this exists. One of them is the percentage of pixels belonging to homogeneous image regions [26]. Meanwhile, noise type and characteristics might influence the estimates of this parameter. Recently, it has been shown that CR obtained for large degrees of image lossy compression might be good indicators of image complexity. To prove this, **Fig. 1.1** shows

dependencies of CR for compression of five typical test color image contaminated by AWGN with two values of the noise variance by the aforementioned BPG coder (mode 4:4:4) (noise is independent in image components). As seen, compression is practically near-lossless for $Q < 29$ (**Fig. 1.1, *a***) and $Q < 33$ (**Fig. 1.1, *b***) due to the influence of the noise for all images. Then, for larger Q, CR starts to grow quickly and for $Q \approx 45$ reaches tens and hundreds depending on image complexity irrespectively to noise variance. CR is the largest for the image Frisco ($E = 5.8$) and is the smallest for highly textural (complex structure) image Baboon ($E = 7.36$). Thus, it seems that CR for fixed PCC corresponding to large degree of compression might serve as an indicator of complexity of noisy images although this should be checked for different models of the noise that might be present in images subject to lossy compression. It is possible to note that similar effects have been observed in compressing grayscale images [27] (see **Fig. 1.2, *a***, the dependence is presented in logarithmic scale) and for the color images using BPG coder (mode 4:2:2, **Fig. 1.2, *b***).

Table 1.1 Four test grayscale images and entropy and MSE values for them ($Q = 20$ and $Q = 25$)

Test image 1 (Legs)	Test image 2 (Fr01)	Test image 3 (Fr02)	Test image 4 (Frisco)
			
$Q = 20$			
$E = 6.50$	$E = 7.46$	$E = 7.40$	$E = 5.82$
MSE = 1.76	MSE = 1.90	MSE = 1.85	MSE = 1.17
$Q = 25$			
MSE = 4.57	MSE = 6.02	MSE = 5.78	MSE = 2.57

Below, let's mainly concentrate on images supposed noise-free. In fact, noise is present in practically all real-life images. However, it is possible to assume an image to be noise-free if noise is invisible. Noise invisibility takes place under certain conditions, e.g., is an image fully textural or has relatively large quasi-homogeneous regions. The fact is that texture masks noise, hence the noise can be noticed in quasi-homogeneous regions if it is intensive enough. For pure additive noise, it can be visible in 8-bit images for white noise if its variance exceeds 20–25 and for spatially

correlated noise if its variance exceeds 6–10. Thus, it is possible to assume an image noise-free if the noise is less intensive than the aforementioned thresholds.

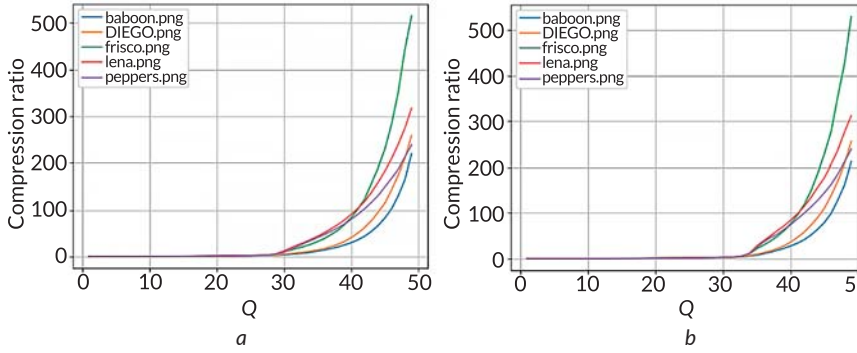


Fig. 1.1 Dependencies of CR on Q for five noisy color images for noise variance equal to 64 (*a*) and 196 (*b*)

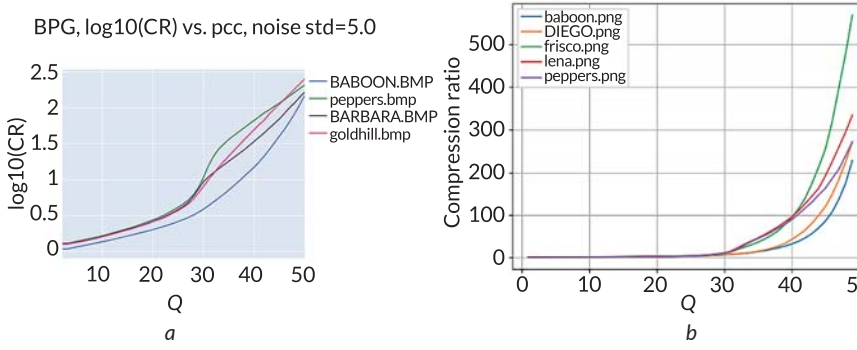


Fig. 1.2 Dependencies of CR on Q for four grayscale images for noise variance equal to 25 (*a*) and five color images for noise variance equal to 100 (*b*) for the BPG coder

Fig. 1.3 gives an example. The test grayscale image Fr03 is corrupted by AWGN with noise variance equal to 25. This noise can be hardly noticed in quasi-homogeneous region in the central part of this image. Meanwhile, the noise is invisible in textural regions, e.g., in the left lower corner of this image, due to the masking effect of texture. This image as well as two images (Fr01 and Fr02) in the central part of **Table 1.1** can be treated as typical examples of images acquired by sensors installed on-board of unmanned aerial vehicles (UAVs) and drones.

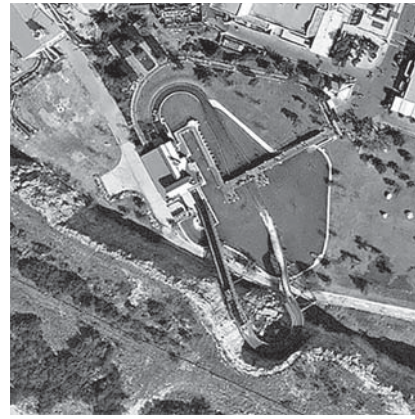


Fig. 1.3 An example of noisy image with hardly noticeable additive white Gaussian noise

Note that, with gaining the popularity of using such sensors, there is practically no difference nowadays between images obtained by remote sensing means (i.e., acquired from some airborne carrier) and standard customer cameras (for some landscapes or urban areas from windows or balconies of multi-store buildings).

1.3 Considered metrics and their properties

If to deal with lossy compression, the introduced distortions have to be characterized quantitatively. For many years, MSE and peak signal-to-noise ratio (PSNR), for single-channel 8-bit image it is determined as PSNR = $10 \log_{10}(255^2/\text{MSE})$ have been the most widely used. In fact, they are still widely used, but the tendency for using visual quality metrics becomes more obvious. There are several reasons behind this. First, as mentioned above, (compressed) images are most often subject to visual inspection and analysis [28–31]. Second, the theory of visual quality metrics has greatly advanced [32–34] in recent 30 years starting from the discrete cosine transform (DCT) based metric DCTune [35] and completing with modern neural network based metrics [36] whilst problems of PSNR in reliable assessment of image visual quality have been shown many times. Third, there is correlation between classification accuracy and both traditional (such as PSNR) and visual quality metrics where the latter ones better correlate with classification accuracy for such "heterogeneous" classes as "Urban" or "Vegetation" in opposite to "homogeneous" classes as "Water surface" or "Grass" [37].

The problems of PSNR can be shown for visually lossless compression that can be also treated as determination of just noticeable difference (JND) point No. 1, i.e. minimal distortions that can be noticed by observers. For example, for the most known coder (standard) JPEG, quality factor (QF) is employed as PCC where QF is integer from 1 to 100 and 100 corresponds to perfect quality (practical absence of distortions). According to data in [9], JND No. 1 for different images takes place for PSNR in the limits from 23 dB to 43 dB and for QF in the limits from 31 to 79. This means the following. First, if one sets PSNR = 43 dB or QF = 79 to guarantee that invisibility of distortions is provided, then for a large percentage of images there is a large space to set a smaller PSNR or QF in order to have a larger CR with still invisible distortions. Second, JND No. 1 is very individual and depends on image complexity.

According to the visual quality metric PSNR-HVS-M [38] (HVS relates to human vision system and M relates to masking) also expressed in dB, the limits of its variation are narrower – from 37.5 dB to 50.5 dB. Thus, it is more adequate but is still not perfect. The situation is similar for some other visual quality metrics [39].

It is difficult to say how many visual quality metrics have been proposed so far. Meanwhile, it is clear that their number exceeds 150 and continues to grow each year. There are several reasons for such a large number of visual quality metrics and attempts to design new ones. First, even the best visual quality metrics are not perfect (this especially relates to no-reference metrics). Second, there is a desire to design universal visual quality metrics [40], i.e. such metrics that are adequate enough for a wide set of distortion types (e.g., there are 24 types of distortions in the database TID2013 [41], here TID2013 relates to Tampere International Database created in 2013). This is a hard task since it is difficult to take into account all peculiarities of human vision system (HVS) and its reaction to specific features of various distortions. The attempts to make this on basis of neural networks that either combine a set of features extracted from an image to be compressed [36, 42] or a set of good visual quality metrics [40, 43, 44] lead to the better universality by the expense of more computations (and time) needed for the metric calculation. Meanwhile, the needed computation time (expanses) is often an important characteristic of a visual quality metric determining its use in practice [45].

While efforts to design better universal metrics continue, a more specific direction of research is to find proper metrics for a particular application such as lossy compression of images in our case. In this sense, the obtained achievements are more obvious. But before starting to consider them, it is necessary to recall how visual quality metrics are compared and analyzed. Usually, image databases with different types of distortions and available assessments (estimates) of image visual quality presented as mean opinion score (MOS) are employed for this purpose. A good visual quality

metric should have a high (at least, rank) correlation between its values and MOS for one or several databases containing images with distortions of interest (saying "high", let's mean that Spearman rank order correlation coefficient (SROCC) has to approach either 1 or –1; the former holds for most metrics that have larger values for better quality, the latter relates to some metrics that have smaller values for better quality such as MSE and Mean Deviation Similarity Index (MDSI) [46]).

Analysis of references dealing with analysis and comparison of metric properties shows the following. Whilst most metrics have SROCC not exceeding 0.9 for all types of distortions in the database TID2013 [43], there are several metrics that have SROCC exceeding 0.92 for the most important distortion types aggregated in subsets "Noise" and "Actual" that correspond to different types of the noise and denoising and lossy compression applications. Recall that PSNR and widely used structural similarity index measure (SSIM) have SROCC values for these subsets smaller than 0.83 and 0.79, respectively.

Moreover, if one considers only distortions dealing with lossy compression, SROCC reaches 0.96–0.97 for such metrics as PSNR-HVS-M [38], MDSI [46], HaarPSI (Haar wavelet-based perceptual similarity index) [47], FSIM (Feature Similarity Index) [48], and some others. Let's note that codes for calculation of aforementioned metrics are available for PSNR-HVS-M, and in Python for MDSI, for HaarPSI, and for FSIM.

Furthermore, there is a very high correlation between the best visual quality metrics [49]. For example, SROCC between MDSI and PSNR-HVS-M is larger than 0.96, between MDSI and HaarPSI – larger than 0.97, between MDSI and FSIM – larger than 0.98 (other data can be found in [49]). And this is not surprising since metrics that have high correlation with MOS should have high correlation between each other. Moreover, good metrics can be recalculated to each other [45]. To show this, **Fig. 1.4** shows the scatter-plot PSNR-HVS-M vs HaarPSI for images with three types of distortions dealing with lossy compression in TID2013. The following expression is valid

$$HaarPSI = -0.0005338 \times PSNR\text{-}HVS\text{-}M^2 + 0.053923 \times PSNR\text{-}HVS\text{-}M - 0.3612, \quad (1.1)$$

where goodness-of-the-fit indicator $R^2 = 0.977$, i.e. is very large.

This means that, for characterizing visual quality of lossy compressed images, one can use any of aforementioned metrics as well as some other good metrics mentioned in [49]. While choosing them for a particular application or analysis, the following properties can be taken into account:

1) experience of working with a given metric including knowledge on approximate positions of JND No. 1 and other JND points;

2) linearity of dependencies of a metric on PCC for a given coder and on MOS for a given database;

3) computational efficiency of metric calculation, etc.

The importance of these properties will become clear later.

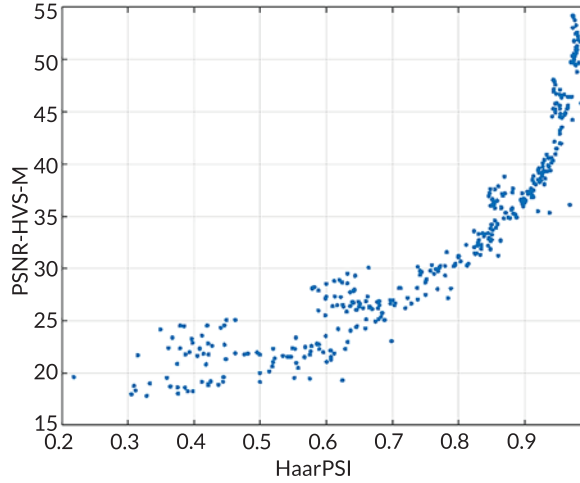


Fig. 1.4 The scatter-plot of PSNR-HVS-M vs HaarPSI for images with three types of distortions dealing with lossy compression in TID2013

Summarizing the brief analysis of metrics able to characterize quality of compressed images, it is possible to state the following. There are several visual quality metrics that are, certainly, not perfect but able to describe quality of lossy compressed images very adequately. To avoid arguments that some compression techniques are based on DCT and a used metric is also based on DCT (similarly for coders based on wavelets and metrics employing wavelets), it is possible to recommend using joint analysis of two metrics based on different transforms or principles of their calculation in studying the performance of lossy compression approaches.

1.4 Some analyzed coders

It is impossible to consider all existing methods of lossy compression. Because of this, let's further rely on data obtained for several typical representatives of compression techniques based on orthogonal transforms since just wavelets and

DCT serve as the basis of most modern practically used coders employed for lossy compression of images and video. The main advantage of this group of methods is that they are quite efficient in different senses – providing quite high quality and ease of variation of compression parameters (CR and quality). Alongside with JPEG, QF is also applied in AV1 Image File Format (AVIF) and High Efficiency Image File Format (HEIF) coders [50, 51]. Bits per pixel (BPP) serves as PCC for JPEG2000 [52] and Set partitioning in hierarchical trees (SPIHT) [53] coders, quantization step (QS) is used in DCT-based coders AGU [54] and advanced DCT (ADCT) [55] coders, scaling factor (SF) serves as PCC in visual quality oriented versions of AGU and ADCT [13] (AGU-M and ADCT-M where M relates to Modified), parameter Q is applied in the BPG coder [27].

Some aspects of using these PCCs will become clear below. They are quite different. JPEG2000 is oriented on easy providing a desired CR, others are more intended on providing a desired quality although, as mentioned above, there are problems with establishing a direct dependence between a PCC and compressed image quality. Let's demonstrate this by two examples. **Fig. 1.5** shows the values of PSNR-HVS-M for the coder AGU-M for two values of BPP – 0.75 (CR = 10.5) and 1.6 (CR = 5).

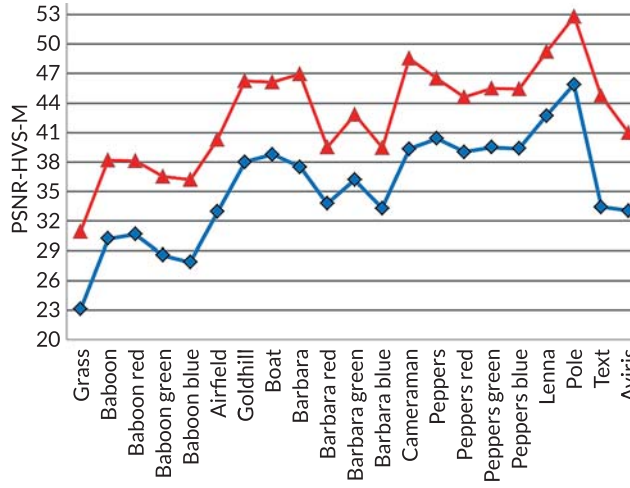


Fig. 1.5 PSNR-HVS-M for several grayscale (component) images compressed by AGU-M with providing two values of BPP, 0.75 (blue color) and 1.6 (red color)

The metric is calculated for images of different origin: conventional grayscale test images, component images of conventional test color images, simple and complex structure test images Pole and Grass, respectively, component image of

AVIRIS (Airborne Visible/InfraRed Imaging Spectrometer) hyperspectral data. As expected, PSNR-HVS-M for BPP = 1.6 is always larger than for BPP = 0.75. Meanwhile, in both cases, PSNR-HVS-M varies in very wide limits – from 23 dB to 46 dB for BPP = 0.75 and from 30 dB to 53 dB for BPP = 1.6. Keeping in mind that PSNR-HVS-M for JND No. 1 is from 37.5 dB to 50.5 dB, it occurs that for some images the distortions are invisible for sure, for some other images it is guaranteed that they are visible whilst for many images (especially for BPP = 1.6) the visibility of distortions is of question. There is also one more observation. The metric values for component images of the same color image (e.g., Barbara) are very close. Let's recall that component images for a given color image are usually highly correlated [56], i.e., similar to each other, and, thus, images of similar complexity are compressed with providing similar quality. It is also seen that if CR increases (BPP decreases) by about 2 times, PSNR-HVS-M diminishes by about 7 dB.

One important question that usually arises in lossy compression is what coder to use? Although there are numerous papers and books showing advantages of newly designed lossy compression techniques, the answer is not simple. A really good coder should outperform its counterparts for different images, different compression ratios, and according to different criteria (metrics). However, this does not hold for any existing coder. **Fig. 1.6** presents dependencies of PSNR-HVS-M on CR for several known and modern compression techniques.

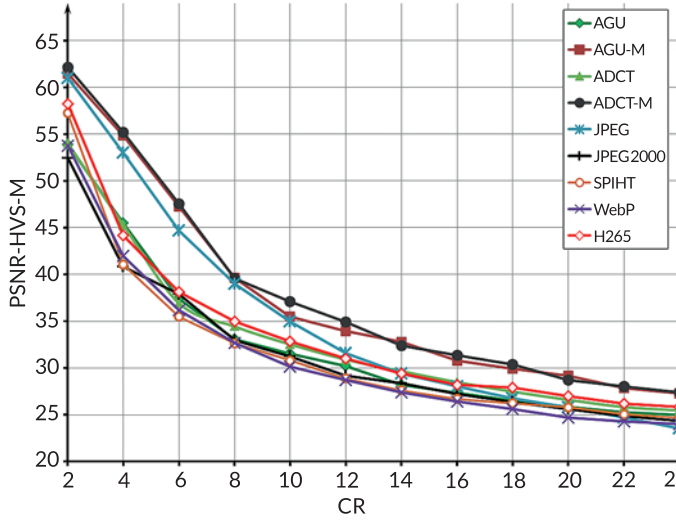


Fig. 1.6 PSNR-HVS-M vs CR for different coders applied to grayscale image Baboon

For *CR* ≈ 4, distortions are invisible for JPEG, ADCT-M, and AGU-M, i.e. coders "adjusted" to peculiarities of HVS by using specific quantization tables for DCT coefficients. For CR ≈ 6 and 8, advantages of these coders compared to other ones remain obvious where distortions are, most probably, visible for CR ≈ 8 for JPEG2000, WebP, AGU, ADCT, H.265, and SPIHT. For CR > 12, JPEG starts to be less efficient than AGU-M and ADCT-M but it performs practically at the same level as other coders. The main reasons behind these observations are that Baboon is a highly textural image and some coders (e.g., H.265 or JPEG2000) are more oriented on larger PSNR than better visual quality.

It might seem that data in **Fig. 1.6** are a particular example. However, there are other data showing that modern coders can perform worse than JPEG for a range of practical situations. For example, for visually lossless compression in JND No. 1 neighborhood, the BPG coder provides twice larger CR on the average than JPEG but the situations when CR for the BPG coder is slightly smaller than for JPEG for the same visual quality are also possible for particular images that usually have complex (textural) structure. In this case, the block size 8 × 8 pixels used in JPEG seems to be the optimal or quasi-optimal solution.

1.5 Rate/distortion curves and their properties

Fig. 1.6 presents an example of the so-called rate/distortion curves (RDCs) that are widely used in coder performance analysis and comparisons. In general, coders performance can be compared in different ways. One of them is to fix CR (or BPP) for all analyzed coders and compare the considered metric values. Another approach presumes fixing the quality metric and compare the CR values where then it is supposed that the best is the coder producing the largest CR. However, even the data in **Fig. 1.6** show the shortcomings of these approaches. Really, JPEG is among the best for CR = 4 whilst it is among the worst for CR = 24. In fact, the best coder should provide the best metric values for a wide (practically important) limits of CR variation, for all images, and for many metrics including PSNR and the best visual quality metrics. According to our experience, this never happens. Even if a coder demonstrates very good results for many images, it always happens that an image can be found for which this coder is not the best. For example, animation or drawings images are very specific. In addition, some coders are designed (optimized) to provide high performance according to PSNR, but then such coders can be not the best according to human vision system (HVS) metrics and vice versa.

Rate/distortion curves can be also used for analysis of compression characteristics depending on image (or image and noise) characteristics. Let's present and consider some examples. They are given in **Fig. 1.7** and **1.8**. Dependencies of PSNR and PSNR-based metrics on quantization step for the AGU and ADCT coders are presented in **Fig. 1.7** for a set of grayscale test images.

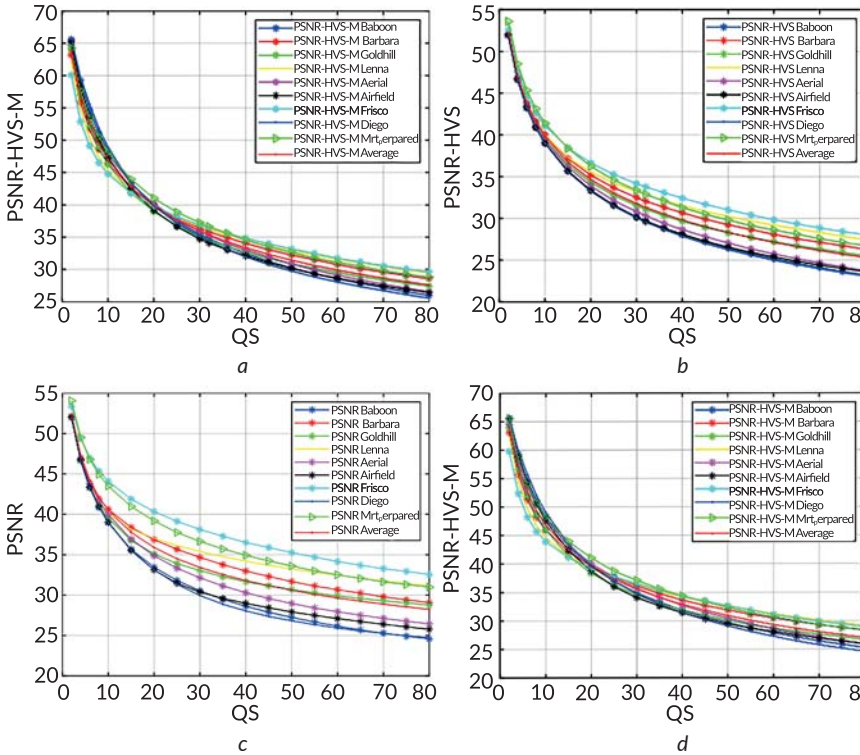


Fig. 1.7 Rate distortion curves for particular test grayscale images: *a* – PSNR-HVS-M(QS) for the AGU coder; *b* – PSNR-HVS(QS) for the AGU coder; *c* – PSNR(QS) for the ADCT coder; *d* – PSNR-HVS-M(QS) for the ADCT coder

The main observations are the following. First, all dependencies are monotonously decreasing. Second, most of them do not cross each other. In other words, if quality (according to a considered metric) of image1 is better than quality of image2 for QS_1, then, most probably, the same holds for other QS. Third, for QS > 20, quality of compressed simple structure images (such as Frisco and MRTprepared) is higher

than quality of complex structure images (Baboon, Diego). Thus, again there is the dependence of compression performance on image complexity. Fourth, locally, it can be assumed that individual dependencies are almost parallel to each other. In other words, for a given QS, they have almost the same derivative *dMetr/dQS* approximately equal to $(dMetr/dQS)_{av}$ obtained for average RDC (such RDCs are given in all four plots) where *Metr* is a considered metric. Moreover, there are intervals of QS (e.g., from 30 to 60) for which the dependencies are almost linear. These properties latter will be further used for quality providing.

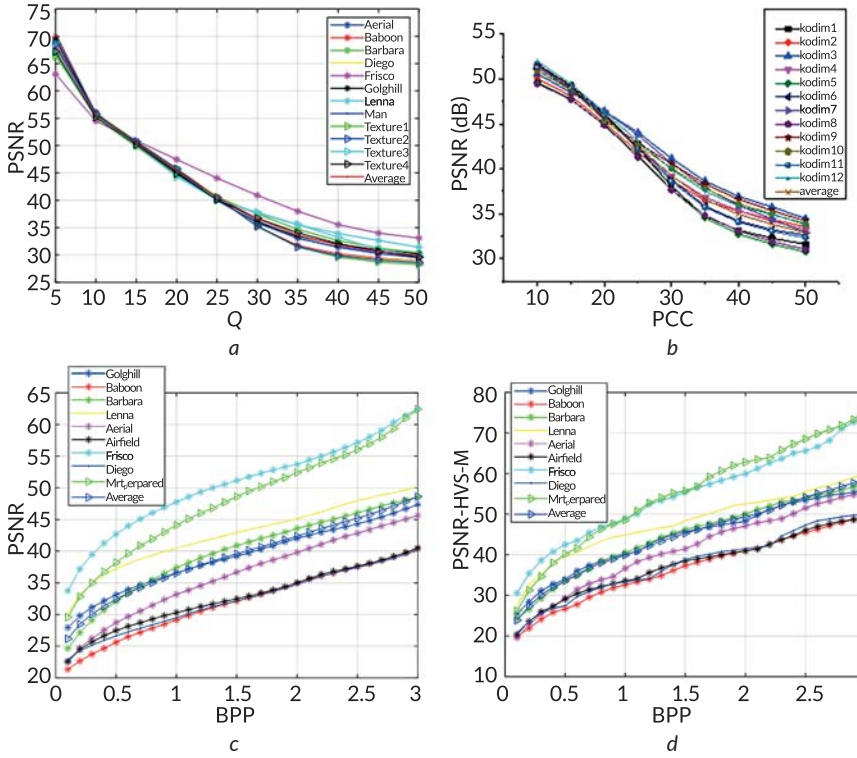


Fig. 1.8 Rate distortion curves for particular test images: *a* – PSNR(Q) for grayscale images compressed by the BPG coder; *b* – PSNR(Q) for color images compressed by the BPG coder; *c* – PSNR(BPP) for SPIHT; *d* – PSNR-HVS-M(BPP) for SPIHT

Analysis of data in **Fig. 1.8** shows the following. The RDCs PSNR(Q) for grayscale images behave in a compact way where only for the test image Frisco the RDC slightly

differs from others. In addition, the dependencies behave almost linearly for a wide range of Q that is of the main interest (from 10 to 35, i.e. from near-lossless to visually lossless compression). Analysis for color images (**Fig. 1.8, *b***) shows that the difference in PSNR for the same Q can be up to 5 dB but there is almost linear (and almost parallel) behavior of RDCs for a wide range of Q variation.

On the contrary, RDCs for SPIHT (**Fig. 1.8, *c*, *d***) have other properties. One reason is that these are dependent on BPP. The best quality according to both PSNR and PSNR-HVS-M is observed for the simple structure images Frisco and MRT prepared where the worst quality takes place for the complex structure image Baboon. The main problems for SPIHT and the corresponding RDCs for it are the following. First, for the same BPP, quality differs a lot. For example, PSNR values for $1 < \text{BPP} < 2$ differ by almost 20 dB. The same holds for the metric PSNR-HVS-M. Second, the derivatives of the metric on BPP are positive and they differ for particular images more than for previously considered RDCs. Although monotonous, the curves PSNR-HVS-M(BPP) have specific "variations".

1.6 Ways of providing a desired quality

First of all, let's explain what can be a desired quality of lossy compressed images. One option is that we would like to provide visually lossless compression. Another option is that, under assumption that a given visual quality metric is able to adequately characterize the image quality, we would like to provide a desired value of a chosen metric. These options might coincide in ways of reaching them but they are not the same. Let's start from the latter option since it can be a part of solving the former task.

Let's recall that it is possible to expect that a chosen metric is in good agreement with subjective estimates of image quality characterized by MOS. MOS values have different scales for different image databases used for image quality assessment. For example, in TID2013, MOS can be potentially from 0 to 9, but, in fact, the minimal value is about 0.2 and the maximal value is 7.2. In the paper [39], it was proposed to divide 3000 distorted images into four groups:

1) images having excellent quality (distortions are invisible or visual quality is considered to be very high); these are images having 200 top rank MOS values;

2) images having good quality; MOS for them have the ranks from 201 to 1000;

3) images of middle quality; for them, MOS have ranks from 1001 to 2000;

4) images of bad quality with the lowest 1000 ranks (from 2001 to 3000).

Note that ranking is carried out in the order of descending MOS. Quantitatively, the images of the first group have MOS larger than 6.05. For the second group,

MOS is in the limits from 5.25 to 6.05. MOS for the third group images is from 3.94 to 5.25. Finally, for the fourth group images, MOS is less than 3.94. This allows determining the thresholds for visual quality metrics using the corresponding scatter-plots and metric value ranking. An example is presented in **Fig. 1.9**. Processing of the obtained data shows that the image quality can be considered as excellent or good if HaarPSI exceeds 0.91. Meanwhile, image quality can be treated as bad if HaarPSI is smaller than 0.75. Similar conclusions can be drawn from analysis of data in **Fig. 1.10** where the scatter-plot of MOS vs HaarPSI is presented only for three types of distortions in TID2013 dealing with lossy compression. The results given in [45] show that invisibility of distortions is provided if HaarPSI exceeds 0.95 or, equivalently, if PSNR-HVS-M is larger than 41 dB (see data in **Fig. 1.4**).

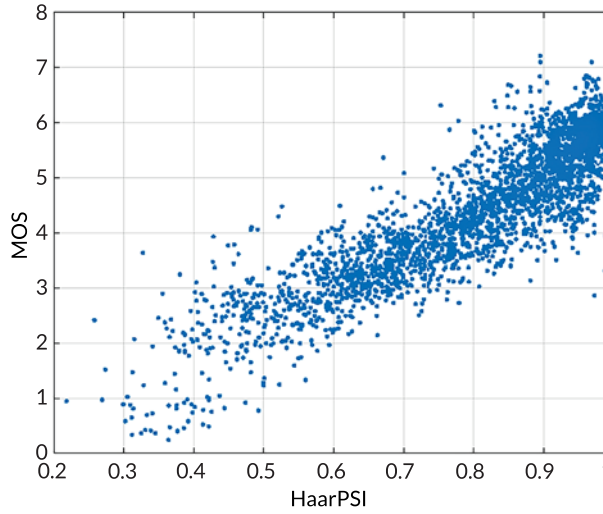


Fig. 1.9 The scatter-plot of MOS vs HaarPSI for all types of distortions in TID2013

It is possible to suppose now that we have decided to characterize image quality by HaarPSI or PSNR-HVS-M equal to 0.91 and 37.5 dB, respectively. It is also possible to assume that we have succeeded to provide them with very high accuracy. This means that MOS varies from approximately 5.2 to 6.0 (**Fig. 1.10**) that corresponds to good quality of compressed images (see above). In turn, MOS equal to 5.6 (the middle of the aforementioned interval) corresponds to HaarPSI from about 0.83 to 0.98 (**Fig. 1.10**). Then, there is no need to provide a desired value of a used metric $Metr_{des}$ with very high accuracy. For example, if $HaarPSI_{des}$ is provided

with errors smaller than 0.01 or the error of providing *PSNR-HVS-M*$_{des}$ is less than 1 dB, this can be treated as acceptable. Detailed analysis has shown that if a given image is compressed by a given coder with two values of PSNR-HVS-M differing from each other by less than 1 dB, they are mostly treated as having equal quality (and often treated as identical).

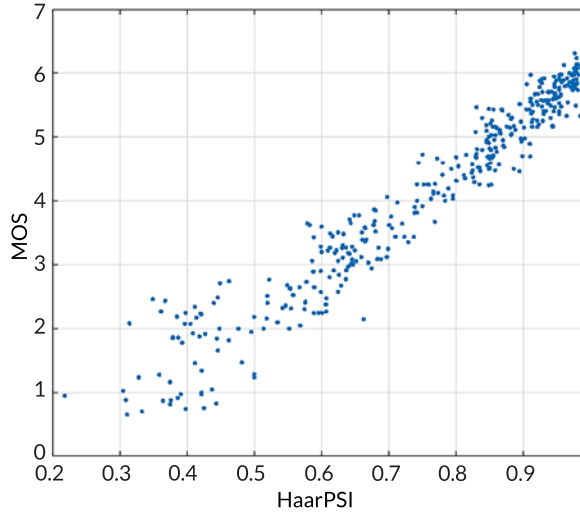


Fig. 1.10 The scatter-plots of MOS vs HaarPSI for three types of distortions dealing with lossy compression in TID2013

Let's now consider the methods of providing a desired quality by reaching some *Metr*$_{des}$. It is possible to start from iterative procedures mentioned above [13, 57, 58]. These methods imply image multiple compression/decompression, metric calculation, and changing PCC to approach *Metr*$_{des}$ with preset accuracy controlled after each iteration. The method performance depends on several factors. First, the starting point (initial value of a used PCC) plays an important role. A way the PCC is changed is important too. At the beginning, PCC can be changed with some discrete step, e.g., ΔQS for AGU or ADCT coders can be set equal to 2. When approaching to vicinity of *Metr*$_{des}$, it is possible to decrease the step of PCC changing or to apply linear interpolation (extrapolation) between the last two values *Metr*(*k*) and *Metr*(*k*–1) where *k* denotes the iteration index. The present accuracy of reaching *Metr*$_{des}$ might have influence too. Finally, for PCC having only integer values (such as QF for JPEG or *Q* for the BPG coder), it can be also so that it is impossible to ensure the desired accuracy

of reaching $Metr_{des}$ less than a certain limit. As it follows from data in **Fig 1.8, *a*, *b***, PSNR decreases by about 1 dB if Q increases by unity for the BPG coder. This means that it is principally impossible to provide $PSNR_{des}$ with error less than 0.5 dB. Some restrictions on accuracy of providing $HaarPSI_{des}$ for this coder are also studied in [45].

Thus, the number of iterations is unpredictable in advance and can be large enough. The results in [13, 57, 58] show that the number of iterations can be up to 10. This can cause problems if there are restrictions on compression time and/or if the compression/decompression time for a given coder is larger. For example, this happens for the ADCT coder for which compression takes certain time due to necessity to optimize partition schemes (decompression is faster). The number of iterations can be decreased for particular coders based on a priori knowledge concerning their main properties [13, 58]. In particular, average curves have been presented in [13] (**Fig. 1.11**) that allow setting the starting *QS* for AGU and ADCT coders (**Fig. 1.11, *a***) and *SF* for AGU-M and ADCT-M coders (**Fig. 1.11, *b***) to provide $PSNR\text{-}HVS\text{-}M_{des}$. The data in **Fig. 1.11, *a*** are in good agreement with data in **Fig. 1.6, *a*** although they have been obtained for different sets of test images. Due to better setting the starting point, it has occurred to decrease the number of iterations by 1.5–3 times [13, 58] with the same accuracy of providing $PSNR_{des}$ and $PSNR\text{-}HVS\text{-}M_{des}$. Moreover, sometimes, it is enough to have only one iteration. However, there are still practical situations where the corresponding computation time can be still inappropriate due to necessity to carry out, e.g., four iterations.

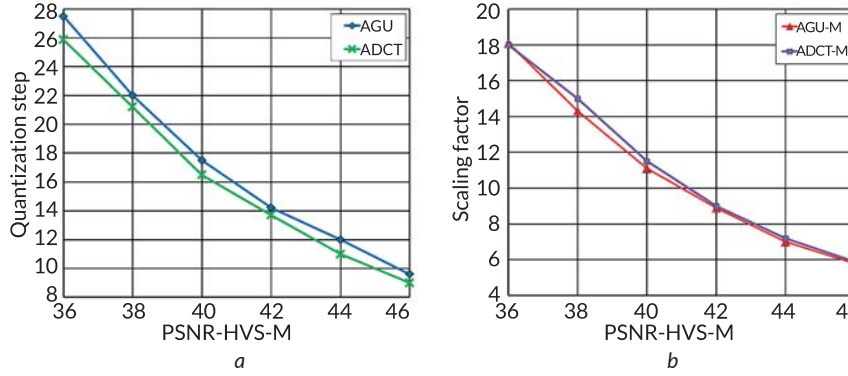


Fig. 1.11 Average curves of QS (*a*) and *SF* (*b*) on PSNR-HVS-M for AGU (*a*), ADCT (*a*), AGU-M (*b*), and ADCT-M (*b*) coders

It might seem that the obtained average curves can be used for setting fixed PCC that corresponds to $Metr_{des}$ for the average curve. This way of providing $Metr_{des}$ is very

fast but it is not accurate. Let's give a few examples to prove this. Suppose that one needs to provide $PSNR_{des}$ = 35 dB for the ADCT coder (see **Fig. 1.7, *c***). Then, one has to set QS ≈ 29. However, depending on an image to be compressed, the provided PSNR values are within the interval 30–38.5 dB, i.e. the errors obviously exceed 1 dB. The situation is slightly better for providing *PSNR-HVS-*M_{des} for the AGU (**Fig. 1.7, *a***) or ADCT (**Fig. 1.7, *d***) coders. Suppose *PSNR-HVS-*M_{des} equals to 35 or 40 dB. In this case, errors are less than 2 dB but, anyway, larger than 1 dB. Similar results have been obtained for 512 × 512 pixel fragments of medical images [59].

The situation is even worse for SPIHT and JPEG2000 (**Fig. 1.8**). Suppose $PSNR_{des}$ = 35 dB. Then, one has to set *BPP* = 0.8 (**Fig. 1.8, *a***). But then the provided PSNR is from 28 dB to 47 dB, i.e. errors are very large. As it follows from [45], $HaarPSI_{des}$ cannot be provided by the BPG coder with appropriate accuracy using fixed *Q* as well. Certainly, we have not checked all coders and all metrics but it seems that it is difficult to reach our goal using fixed PCC.

Then, some trade-off is needed. This can be a two-step approach described in detail in the thesis of L. Fangfang [60] and the basic papers [15, 16, 61–64]. The idea is the following. Let's obtain RDCs of interest in advance for a set of test images with various properties and get the average RDC for them with estimation of derivative values for different fragments of this curve. Then, for a given image to be compressed by a given coder with providing a given $Metr_{des}$, let's determine PCC_1 (a starting PCC) that corresponds to $Metr_{des}$ for the average curve. Then, carry out compression and decompression for the considered image using PCC_1 and estimate $Metr_1$ for the decompressed image. If $Metr_1$ does not differ from $Metr_{des}$ essentially (i.e. $|Metr_1 - Metr_{des}| < \delta_M$ where δ_M is the allowed error), then stop and keep the compressed image as the desired result. If no, calculate PCC_2 as

$$PCC_2 = PCC_1 + \left(Metr_{des} - Metr_1\right) \Big/ \left(dMetr\left(PCC_1\right) / dPCC\right), \tag{1.2}$$

where rounding-off should be applied if *PCC* can be only integer.

As seen, linear approximation of the RDCs based on the average curve is applied within this approach. This explains both its advantages and shortcomings. The advantages are twofold. First, the method usually provides about one order less errors (variance) of providing $Metr_{des}$ after the second step than after the first step where improvement is provided almost for all images and in the wide range of $Metr_{des}$ [16, 60, 63]. Second, the procedure of compression with providing a given $Metr_{des}$ usually takes considerably less time than iterative procedures described above since only two compressions and one decompression are needed (and sometimes the second compression is not needed at all). The two-step procedure occurs to be efficient for RDCs presented in **Fig. 1.8, *b***.

The main shortcomings are considered in [65]. First, it might be so that the mean of $Metr_2$ is biased and then one has to take into account not only variance of the provided metric values but the bias as well. This happens if the absolute value of the second derivative of the average curve in PCC_1 is large enough. Till the moment, no attempts to take into account the second derivative have been undertaken. Second, due to the nonlinearity of the average curve, it might be so that the determined PCC_2 occurs to be negative. This is in contradiction with assumptions concerning the range of PCC (for all types of PCC, it can be only positive).

Let's consider such "bad" situations. Suppose that one has to provide $PSNR_{des} = 30$ dB for the AGU coder. Then, QS_1 is set equal to 60.15. There are test images (Goldhill, Barbara) for which $PSNR_{des}$ is provided with appropriate accuracy, i.e. with the errors less than 1 dB) even for QS_1 although the second step provides practically perfect result. Meanwhile, there are images (e.g., Baboon, Frisco, and Diego) for which $PSNR_1$ differs from $PSNR_{des}$ significantly. After the second step, the situation does not radically improve since the errors are still about 3–5 dB for such textural images as Diego and Baboon. The reason for this is large difference QS_1QS_2, i.e. QS_1 and QS_2 differ a lot and linear approximation of the particular and average RDCs is not valid anymore. Finally, there are images (Lenna, Aerial) for which the second step improves accuracy (decreases the errors) of $PSNR_{des}$ providing. Due to such images as Diego and Baboon, variance after the second step does not decrease significantly compared to variance after the first step (**Table 1.2**).

Table 1.2 Statistics and parameters of providing $PSNR_{des} = 30$ dB for a set of test images

Test image	QS_1	$PSNR_1$	ΔQS	QS_2	$PSNR_{prov}$
Goldhill	60.15	30.34	4.24	64.39	29.98
Baboon	60.15	26.72	–40.2	19.87	33.97
Barbara	60.15	30.88	10.87	71.01	29.95
Lenna	60.15	32.62	32.35	92.50	30.72
Aerial	60,15	28.34	–20.4	39.68	30.69
Airfield	60.15	27.05	–36.3	23.80	32.52
Frisco	60.15	34.68	57.62	117.7	31.16
Diego	60.15	26.44	–43.83	16.31	35.21
$Mrt_{\_prepared}$	60.15	32.89	35.66	95.81	24.51
Variance		9.12			9.05

Then, special restrictions can be imposed [66]. For example, PCC_2 is set equal to $PCC_1/2$ if the calculated PCC_2 is less than $PCC_1/2$. This does not allow PCC_2 to be negative (since PCC_1 is always positive) and, in fact, restricts the area where linear

approximation is considered appropriate. Third, there are cases (RDCs for particular images) where dependences *Metr*(*PCC*) and $dMetr(PCC_1)/dPCC$ for a given image to be compressed differs from the average curve and derivative a lot (see **Fig. 1.8, *c*, *d***). Then, although the two-step procedure improves the accuracy considerably, this accuracy is still inappropriate [67]. A way out is proposed in [61] where it is taken into account that the DCT-based coder AGU provides almost the same compression characteristics as SPIHT and CR for the AGU coder can be easily predicted. One more solution to improve accuracy of providing $Metr_{des}$ has been proposed in [62, 64]. It has been taken into account that RDCs usually depend on image complexity. Then, it is possible to obtain average curves for simple and complex structure images in advance. For a given image to be compressed, its complexity is determined first, and after this, based on such a pre-classification, the corresponding average RDC is used within the two-step procedure to determine PCC_2.

Although the trade-off had been found, the desire was still to design an approach to adaptive compression without decompression and the second step of compression. Because of this, in parallel with the design of the two-step approach, the approaches based on prediction have been studied [10, 68–70]. The idea put behind the approach proposed in [68] is that, for JPEG with uniform quantization, MSE of distortions introduced due to quantizing DCT coefficients in blocks is equal to MSE of distortions observed in the corresponding decompressed image blocks, i.e., in spatial domain. Then, it is possible to calculate aggregate MSE_{DCT} in the DCT domain for a given *QS* and estimate (predict) MSE_{spat} of distortions after decompression (due to compression). MSE in DCT domain with averaging the local estimates for all blocks can be expressed as

$$MSE_{DCT} = \frac{1}{63N}\sum_{n=1}^{N}\left(\sum_{i=1}^{63}\left(QS \times round\left(\frac{D_i}{QS}\right) - D_i\right)^2\right), \tag{1.3}$$

where D_i is the value of the i-th DCT coefficient in a processed n-th block, N denotes the number of analyzed blocks, 63 AC (alternating current) DCT coefficients are analyzed, round denotes rounding-off to the nearest integer. The value MSE_{spat} has been shown to be approximately equal to MSE_{DCT}. In addition, it has been demonstrated that there is no need to take all possible 8 × 8 pixel blocks to calculate MSE_{DCT}. It was enough to have 500 blocks. Since DCT in 8 × 8 pixel blocks is a fast operation, prediction for a given *QS* is very fast. Moreover, it has been shown [68], that, using correcting factors with values close to unity, it is possible to predict MSE for given *QS* for the AGU and ADCT coders. The prediction accuracy was shown to be better than for the procedure in [71].

The paper [69] shows that the same prediction can be carried out for non-uniform quantization often used in JPEG and employed in the AGU-M and ADCT-M coders. For the latter two coders that, in fact, use 32 × 32 pixel and adaptive block size, respectively, the prediction based on 8 × 8 pixel blocks is shown possible with correction factors of about 0.9 and relative error less than 10 %, i.e. smaller than 0.4 dB which is appropriate for practice. The prediction approach is tested in [69] for a wide set of images of different origin.

The paper [70] proposes another approach to prediction. It is shown [10, 12, 70] that for small *QS* for DCT-based coders the *MSE* of introduced distortions is often proportional to QS^2 and approximately equal to $QS^2/12$. For larger *QS*, the dependence becomes less fast than ~QS^2 and depends on image complexity. Then, it is proposed to predict *MSE* as

$$MSE \approx \left(QS^2/12\right) f(X), \tag{1.4}$$

where $f(X)$ is some function of one or several parameters describing image complexity. As one option, the use of probability P_0 that DCT coefficients in 8 × 8 pixel blocks occur to be zero after quantization using a given *QS* is considered. This probability is small for small *QS* and it increases if *QS* becomes larger. Note that for $P_0 < 0.6$ the function $f(X)$ slowly decreases but it is almost equal to unity. Decreasing speed increases if P_0 exceeds 0.6. The approximations are given in [70] and for $P_0 \approx 0.9$ one has $f(X) \approx 0.6$, but the accuracy is not good enough.

Let's note that the approach [70] is based on obtaining the scatter-plot for a set of test images compressed with different *QS*. Slightly other approaches are studied in [12, 72]. In [72], it is shown that there is an essential dependence between image entropy *E* (that deals with image complexity of noise-free images) and MSE of introduced distortions for the BPG coder applied to grayscale images. One example from the paper [72] is presented in **Fig. 1.12** where MSE values are given for 21 test images compressed with $Q = 25$. The obtained results show the following: can vary from about 4 to 8 where for most images *E* exceeds 7.0 and the MSE values about 6 are observed for them. Meanwhile, there are also quite many images having smaller entropy and MSE values for them are considerably smaller (except the test image Goldhill). This example, as well as other examples given in [72], show that there is a considerable correlation between MSE and *E*. Further studies [74] have shown that MSE has a certain degree of correlation with other parameters characterizing complexity and the joint use of such parameters can improve prediction where PCC and a few parameters characterizing complexity are jointly used as inputs of a very simple trained neural network.

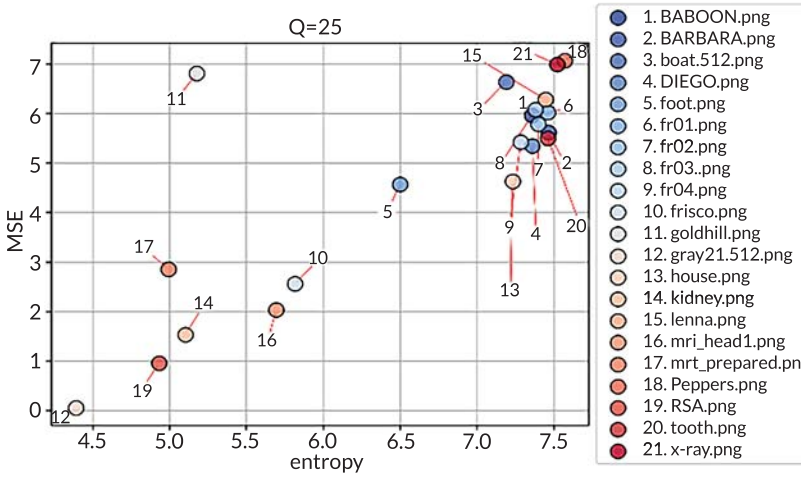
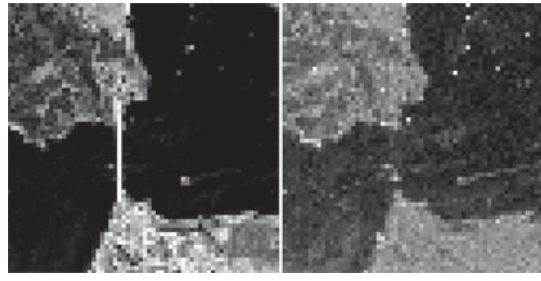


Fig. 1.12 Scatter-plot of MSE vs *E* for grayscale images compressed by the BPG coder

The data in [72, 73] partially explain why MSE of introduced distortions is usually larger for more complex structure images. A detailed local analysis of distortion statistics is carried out and it is shown that distortions are larger in image regions (blocks) having higher activity characterized by local variance. Then, if an image has more blocks identified as locally active, integral MSE should be larger. To illustrate this phenomenon, **Fig. 1.13** presents two maps for the test image Frisco (**Table 1.1**, the rightmost image): the map of local activity (**Fig. 1.13, *a***, more active blocks are shown by lighter color) and the map of local MSEs of distortions (**Fig. 1.13, *b***, larger MSEs are indicated by brighter color). As seen, there is a high "correlation" between these two maps.



a *b*

Fig. 1.13 The local activity map (*a*) and the map of local MSEs of distortions due to lossy compression (*b*)

Currently, the accuracy of MSE prediction by the methods described above is at the same order as for the two-step methods. The task dealing with MSE prediction is that the prediction does not directly solve the problem of providing MSE_{des}. Then, one has to start from some PCC_{start} and predict MSE_{start} for it. If this MSE_{start} differs from MSE_{des} considerably (more than allowed), PCC has to be changed accordingly (one way to do this is described in [72]). Then, the processing becomes two-step but, recall, there is no need to compress and decompress the image at the first step. Thus, the compression with providing MSE_{des} can be faster than for two-step method described above if prediction requires less time expenses than compression and decompression. This is (could be) the main advantage of the approaches based on prediction. However, there are also several drawbacks. First, currently, the prediction approach has been mainly oriented on determining MSE or, equivalently, PSNR. Then, it is needed to check whether or not this approach allows predicting visual quality metrics. Second, the trade-off between complexity and accuracy of prediction has to be found for coders under interest. Third, fast and accurate algorithms of setting PCC for final compression have to be designed and tested.

Finally, let's consider the case of visually lossless compression based on JND No. 1 [9, 75]. This approach presumes that, depending on an image property indicator, a metric or PCC value ($Metr_{JND}$ or PCC_{JND}) that corresponds to JND No. 1 is determined and, then, this value is provided. The approach is based on preliminary analysis for special databases (such as KonJND-1k [76] and some others) that contain many images for which JND No. 1 points have been estimated for a considered coder by experiment participants.

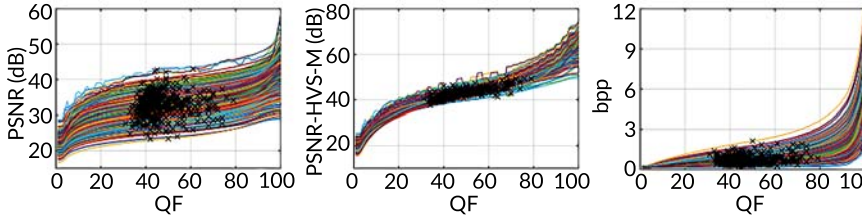
Some results of experiments for JPEG are demonstrated in **Fig. 1.14** taken from [9]. JND No. 1 point for each RDC is illustrated by ×. The following conclusions can be drawn:

1) JND No. 1 points are observed for a wide range of QF and PSNR values (**Fig. 1.14, *a***), for a narrower range of PSNR-HVS-M values (**Fig. 1.14, *b***), and a wide range of bpp values (**Fig. 1.14, *c***); there are some (rarely met) images for which even the use of QF ≈ 80 does not guarantee invisibility of distortions whilst for some images the distortions are invisible for QF ≈ 30;

2) there are some curves PSNR(QF) and PSNR-HVS-M(QF) that are not monotonous corresponding to the so-called strange images [14] for which many methods of providing $Metr_{des}$ might have problems.

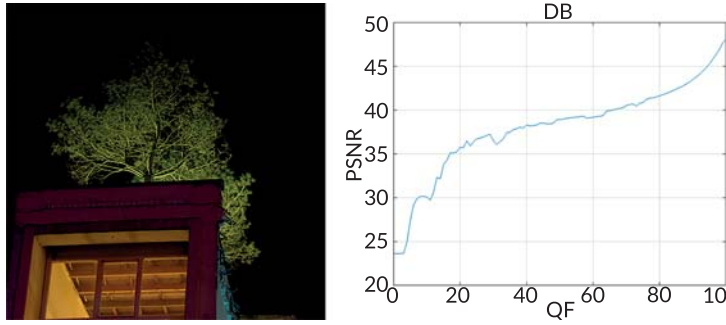
Let's give a few examples of strange images and RDCs for them. **Fig. 1.15** presents natural scene color image (**Fig. 1.15, *a***) and RDC PSNR vs QF for JPEG (**Fig. 1.15, *b***). As seen, the image has a large quasi-homogeneous dark region and the curve has jumps with the amplitude reaching 1 dB. An example of remote sensing image is giv-

en in **Fig. 1.16, *a*** and RDCs PSNR vs QS for color and intensity components compressed by the AGU coder are presented in **Fig. 1.16, *b***. Obviously, all four curves have "fluctuations" and are not monotonous.



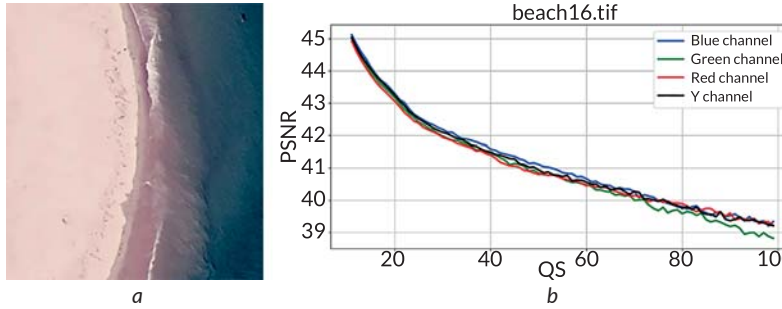
a *b* *c*

Fig. 1.14 RDCs PSNR (*a*), PSNR-HVS-M (*b*), and bpp (*c*) on QF, along with JND No. 1 points, for a subset of scene color images taken from the KonJND-1k database and compressed by JPEG



a *b*

Fig. 1.15 Natural scene color image (*a*) and RDC PSNR vs QF for JPEG (*b*)



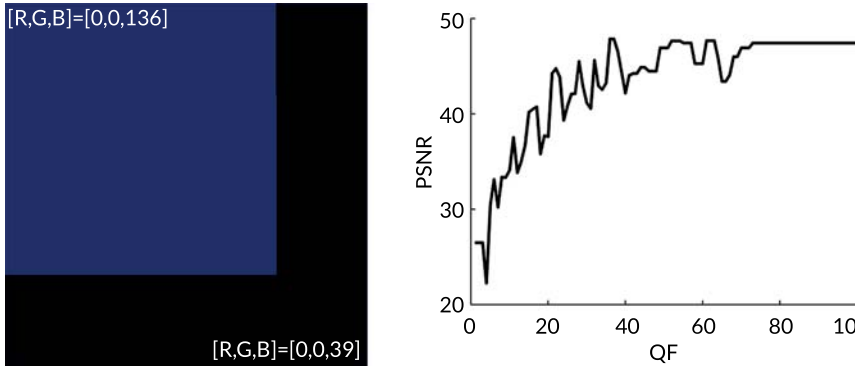
a *b*

Fig. 1.16 Remote sensing image (*a*) and the RDCs PSNR vs QS for AGU (*b*)

Fig. 1.17 shows two other color images that occur to be strange if the dependence PSNR(QF) for JPEG is considered. Joint analysis of images in **Fig. 1.15, *a***, **Fig. 1.16, *a***, and **Fig. 1.17** shows that images can be strange if they contain large dark, or white (e.g., overexposed images) or other color regions. This assumption has been checked for artificial color image presented in **Fig. 1.18, *a*** compressed by JPEG and the assumption has been confirmed (see RDC of *PSNR* vs *QF* in **Fig. 1.18, *b***).

a *b*

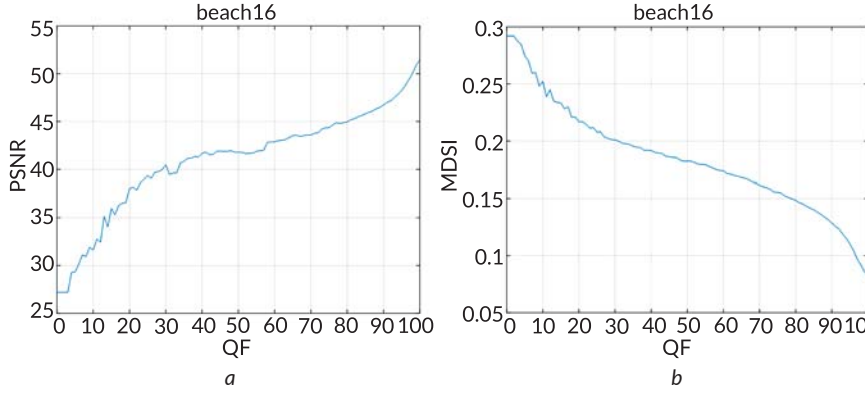
Fig. 1.17 Examples of two strange images: example 1 (*a*) and example 2 (*b*)



a *b*

Fig. 1.18 Artificial color test image with large homogeneous regions (*a*) and RDC PSNR vs QF for JPEG (*b*)

It might seem that strange images take place only according to PSNR and mainly for JPEG. However, this is not true. **Fig. 1.19** shows dependencies of PSNR and MDSI metrics on QF for intensity component of the color image beach16 represented in **Fig. 1.16, *a***. As seen, beach16 is the strange image not only according to PSNR but also according to the visual quality metric MDSI. Note that the largest "fluctuations" of MDSI take place in the same interval of QF as for PSNR – for QF about 10.



a *b*

Fig. 1.19 Dependencies PSNR (*a*) and MDSI metric (*b*) on QF for intensity component of the color image beach16

The images strange for JPEG often occur to be strange for the coder AGU. This is not surprising since both use DCT in blocks as the basis for lossy compression.

In addition, not only color images can be strange – grayscale images can be strange as well especially those ones having large homogeneous backgrounds, e.g., images of text documents or advertising materials.

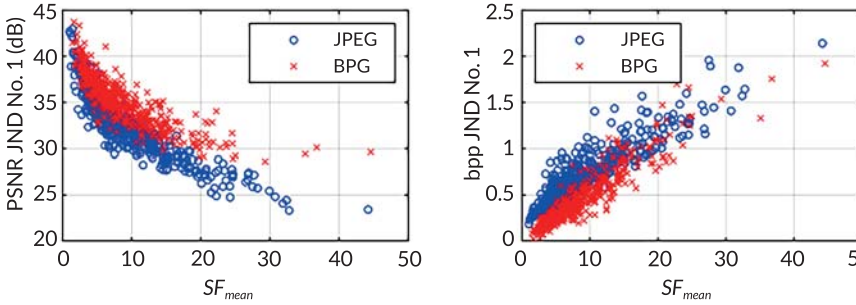
The cases and reasons for strange images are discussed in [14, 77]. Currently, we deal with detection of strange images before compression.

Let's come back to considering JND No. 1. It is shown in [9] that there are many image property (complexity) indicators that are correlated with parameters of JND No. 1. Mean spatial frequency SF_{mean} that can be quite easily calculated has the largest correlation [75]. This can be seen by analysis of scatter-plots in **Fig. 1.20**. Such scatter-plots have allowed to provide the following rules for determining PSNR for JPEG and BPG, respectively:

$$PSNR(SF_{mean}) = \begin{cases} 0.01376SF_{mean}^2 - 0.8645SF_{mean} + 38.97, & SF_{mean} \le 31; \\ 25.4, & SF_{mean} > 31. \end{cases} \quad (1.5)$$

$$PSNR\left(SF_{mean}\right)=\begin{cases}0.01336SF_{mean}^{2}-0.7618SF_{mean}+40.36, & SF_{mean}\le 28;\\ 29.5, & SF_{mean}>28.\end{cases} \quad (1.6)$$

More details can be found in [14, 75, 78, 79]. Thus, either the metric value or PCC that correspond to JND No. 1 can be determined and the next step is to provide it. As shown above, this task can be solved by the two-step or prediction approaches.



a *b*

Fig. 1.20 Scatter-plots of PSNR (*a*) and *bpp* (*b*) vs SF_{mean} for the first JND points determined using KonJND-1k image subsets for JPEG and BPG coders

The JND No. 1 position represents the boundary between visually lossless and visually lossy compression. Therefore, these positions can be used in the assessment and analysis of image quality [80]. Thus, **Fig. 1.21** illustrates how PSNRs that correspond to JND No. 1 can be used to detect high quality images ($PSNR > PSNR_{JND\ No.\ 1}$). In the quality analysis, well-known publicly available image datasets from the visible (LIVE and VCL@FER) and the infrared part of the electromagnetic spectrum (long wave infrared – LWIR [80]) are used. Using $PSNR_{JND\ No.\ 1}$, JPEG compressed images of excellent quality are detected in LIVE (see lower difference MOS (DMOS) values, **Fig. 1.21, *a***) and LWIR datasets (higher MOS values, **Fig. 1.21, *c***), while for the VCL@FER database there are several exceptions with good quality images (**Fig. 1.21, *b***).

The results in [79] demonstrate that the influence of compression at the JND No. 1 does not significantly affect the target detection by humans. This allows compressing images in adaptive way instead of using fixed QF for JPEG with saving memory effectively. Meanwhile, accuracy of approaches based on JND No. 1 needs additional study.

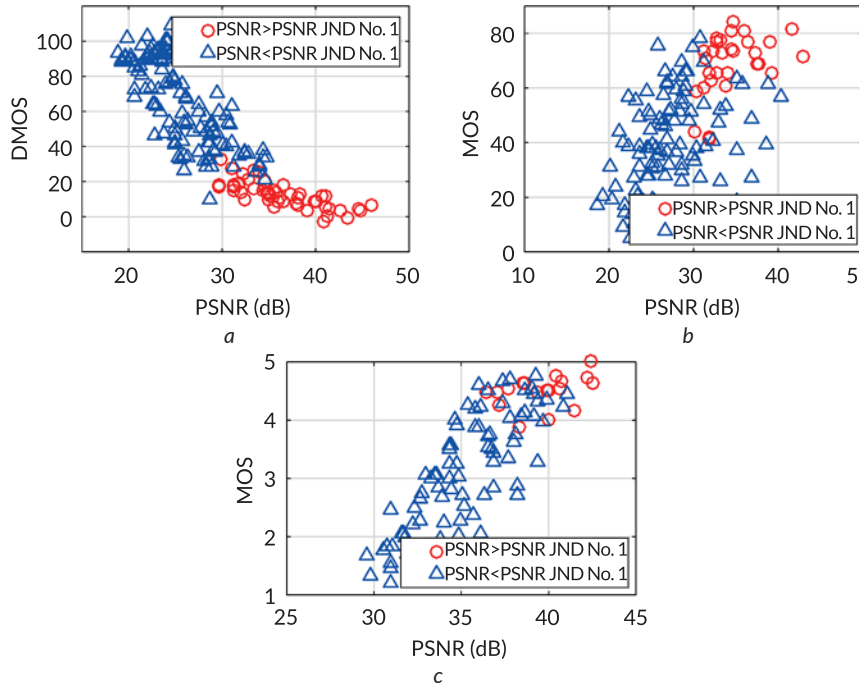


Fig. 1.21 Scatter plots of PSNR and subjective quality scores for JPEG images from three datasets (*a*) LIVE (175 images), (*b*) VCL@FER (138 images) and (*c*) LWIR (100 images)

We would like to discuss some limitations of the considered approaches. First, if one wants to apply the two-step approach for a new metric and/or a new coder, the corresponding average RDC and its derivative have to be obtained. This is not a difficult work and it has to be done once. Some other drawbacks and limitations of the two-step approach are considered in [65].

Second, the method of MSE and MSE_{HVSM} prediction [68–70, 72] has been tested for several DCT-based compression techniques. However, it is not clear can it be adapted to wavelet-based coders. Besides, it is not clear can the approach be applied to prediction of other metrics. The approach to MSE prediction based on indicators of image complexity [12] is at the beginning of its design and it has been tested only for the BPG encoder. It is possible to expect that the approach should be useful for other compression techniques, but to be general, it has to be modified and verified for other coders.

Finally, the approach based on JND [9, 80] presumes the use of several databases of distorted images and the results of their quality assessment by a great number of experiment participants. Such experiments are labor and time consuming and have to be carried out for each compression technique under interest. However, such experiments can be done once and then the obtained results can be used for several purposes.

1.7 Conclusions and future research directions

Summarizing the presented materials, it is possible to conclude the following:

1. There are quite many metrics of visual quality that are able to quite adequately perform for lossy compression of images.

2. In general, there are several ways to provide a desired quality according to the chosen metric; advantages and drawbacks of these ways have been discussed above from the viewpoints of accuracy and time needed for their realization.

3. Some of these ways can be further advanced; this relates to the two-step methods, algorithms based on prediction, and methods relating to JND No. 1.

4. More research should be done for prediction of visual quality metrics, especially for color and multichannel images.

5. Strange images have to be considered more in detail; reliable and fast detection of potential strange images before lossy compression is a task to be solved.

6. Approaches to image quality providing for modern coders such as HEIF and AVIF has to be studied; to the best of our knowledge, the corresponding research has not been carried out yet.

7. Applicability of the considered approaches to providing the quality of infrared images has to be studied as well.

8. It is also worth continuing studies of lossy compression impact on classification and recognition of images and remote sensing data.

Conflict of interest

The authors declare that there is no conflict of interest in relation to this paper, as well as the published research results, including the financial aspects of conducting the research, obtaining and using its results, as well as any non-financial personal relationships.

Financing

The study was performed without financial support.

Data availability

Data will be made available on reasonable request.

Use of artificial intelligence statement

The authors confirm that they did not use artificial intelligence technologies when creating the current work.

Authors' contributions

Volodymyr Lukin: Conceptualization, Supervision, Writing – review & editing.
Sergii Kryvenko: Data curation, Investigation.
Fangfang Li: Methodology, Writing – original draft, Software.
Sergiy Abramov: Methodology, Data curation.
Viktoriia Abramova: Project administration, Data curation.
Bohdan Kovalenko: Resources, Software.
Igor Dohtiev: Software, Formal analysis.
Oleksandr Arkhipov: Software, Formal analysis.
Nenad Stojanović: Formal analysis, Writing – original draft, Software.
Boban Bondžulić: Methodology, Data curation, Resources.

References

1. Wei, J., Mi, L., Hu, Y., Ling, J., Li, Y., Chen, Z. (2022). Effects of Lossy Compression on Remote Sensing Image Classification Based on Convolutional Sparse Coding. IEEE Geoscience and Remote Sensing Letters, 19, 1–5. https://doi.org/10.1109/lgrs.2020.3047789
2. Dougherty, G. (2009). Digital Image Processing for Medical Applications. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511609657

3. Doss, S., Pal, S., Akila, D., Jeyalaksshmi, S., Jabeen, T. N., Suseendran, G. (2020). Satellite image remote sensing for identifying aircraft using SPIHT and NSCT. Journal of Critical Reviews, 7 (5), 631–634. https://doi.org/10.31838/jcr.07.05.130
4. Zabala, A., Pons, X. (2011). Effects of lossy compression on remote sensing image classification of forest areas. International Journal of Applied Earth Observation and Geoinformation, 13 (1), 43–51. https://doi.org/10.1016/j.jag.2010.06.005
5. Radosavljević, M., Brkljač, B., Lugonja, P., Crnojević, V., Trpovski, Ž., Xiong, Z., Vukobratović, D. (2020). Lossy Compression of Multispectral Satellite Images with Application to Crop Thematic Mapping: A HEVC Comparative Study. Remote Sensing, 12 (10), 1590. https://doi.org/10.3390/rs12101590
6. Lyalko, V., Popov, M., Sedlerova, O., Fedorovskyi, O., Stankevich, S., Yelistratova, L. et al. (2022). On the development of remote sensing methods and technologies in Ukraine. Ukrainian Journal of Remote Sensing, 9 (2), 43–53. https://doi.org/10.36023/ujrs.2022.9.2.214
7. Altamimi, A., Ben Youssef, B. (2024). Lossless and Near-Lossless Compression Algorithms for Remotely Sensed Hyperspectral Images. Entropy, 26 (4), 316. https://doi.org/10.3390/e26040316
8. Christophe, E.; Prasad, S., Bruce, L. M., Chanussot, J. (Eds.) (2011). Hyperspectral Data Compression Tradeoff. Optical Remote Sensing. Berlin, Heidelberg: Springer 9–29. https://doi.org/10.1007/978-3-642-14212-3_2
9. Bondžulić, B., Stojanović, N., Lukin, V., Kryvenko, S. (2024). JPEG and BPG visually lossless image compression via KonJND-1k database. Vojnotehnicki Glasnik, 72 (3), 1214–1241. https://doi.org/10.5937/vojtehg72-50300
10. Krivenko, S. S., Krylova, O., Bataeva, E., Lukin, V. V. (2018). Smart lossy compression of images based on distortion prediction. Telecommunications and Radio Engineering, 77 (17), 1535–1554. https://doi.org/10.1615/telecomradeng.v77.i17.40
11. Kovalenko, B., Lukin, V. (2023). Analysis of distortions due to BPG-based lossy compression of noise-free and noisy images. Herald of Khmelnytskyi National University. Technical sciences, 325 (5), 128–135.
12. Kovalenko, B., Lukin, V. (2024). Pre-Requisites for Mean Square Error Prediction in Better Portable Graphics Based Lossy Compression of Grayscale Images. 2024 IEEE 42nd International Conference on Electronics and Nanotechnology (ELNANO). Kyiv, 488–492. https://doi.org/10.1109/elnano63394.2024.10756949
13. Zemliachenko, A., Lukin, V., Ponomarenko, N., Egiazarian, K., Astola, J. (2015). Still image/video frame lossy compression providing a desired visual quality. Multidimensional Systems and Signal Processing, 27 (3), 697–718. https://doi.org/10.1007/s11045-015-0333-8

14. Bondzulic, B., Bujakovic, D., Li, F., Lukin, V. (2022). On strange images with application to lossy image compression. Radioelectronic and Computer Systems, 4, 143–152. https://doi.org/10.32620/reks.2022.4.11
15. Li, F., Lukin, V., Okarma, K., Fu, Y. (2021). Providing a Desired Quality of BPG Compressed Images for FSIM Metric. 2021 IEEE 3rd International Conference on Advanced Trends in Information Theory (ATIT). Kyiv, 10–14. https://doi.org/10.1109/atit54053.2021.9678522
16. Li, F., Lukin, V., Ieremeiev, O., Okarma, K. (2022). Quality Control for the BPG Lossy Compression of Three-Channel Remote Sensing Images. Remote Sensing, 14 (8), 1824. https://doi.org/10.3390/rs14081824
17. Bondžulić, B., Pavlović, B., Stojanović, N., Petrović, V., Bujaković, D. (2023). A simple and reliable approach to providing a visually lossless image compression. The Visual Computer, 40 (5), 3747–3763. https://doi.org/10.1007/s00371-023-03062-y
18. Ponomarenko, N. N., Lukin, V. V., Egiazarian, K. O., Lepisto, L. (2013). Adaptive visually lossless JPEG-based color image compression. Signal, Image and Video Processing, 7 (3), 437–452. https://doi.org/10.1007/s11760-013-0446-1
19. Makarichev, V., Vasilyeva, I., Lukin, V., Vozel, B., Shelestov, A., Kussul, N. (2021). Discrete Atomic Transform-Based Lossy Compression of Three-Channel Remote Sensing Images with Quality Control. Remote Sensing, 14 (1), 125. https://doi.org/10.3390/rs14010125
20. Makarichev, V., Lukin, V., Brysina, I. (2024). On the Impact of Discrete Atomic Compression on Image Classification by Convolutional Neural Networks. Computation, 12 (9), 176. https://doi.org/10.3390/computation12090176
21. Jia, S., Ji, Z., Qian, Y., Shen, L. (2012). Unsupervised Band Selection for Hyperspectral Imagery Classification Without Manual Band Removal. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 5 (2), 531–543. https://doi.org/10.1109/jstars.2012.2187434
22. Ponomarenko, N. N., Lukin, V. V., Zriakhov, M. S., Kaarna, A., Astola, J. (2008). Automatic Approaches to On-Land/On-Board Filtering and Lossy Compression of AVIRIS Images. IGARSS 2008 – 2008 IEEE International Geoscience and Remote Sensing Symposium. Boston, III-254–III–257. https://doi.org/10.1109/igarss.2008.4779331
23. Lukin, V., Krivenko, S., Li, F., Abramov, S., Makarichev, V. (2022). On Image Complexity in Viewpoint of Image Processing Performance. Information technologies and systems of information security. IntelITSIS-2022, 16.
24. Bondžulić, B., Lukin, V., Bujaković, D., Li, F., Kryvenko, S., Pavlović, B. (2023). On Visually Lossless JPEG Image Compression. 2023 Zooming Innovation in

Consumer Technologies Conference (ZINC), 113–118. https://doi.org/10.1109/zinc58345.2023.10174090
25. Zhang, Y., Zhang, Z., Wang, X., Wang, X., Ge, J., Bian, F. (2018). An Adaptive Infrared Image Preprocessing Method Based on Background Complexity Descriptors. 2018 Eighth International Conference on Instrumentation & Measurement, Computer, Communication and Control (IMCCC), 344–349. https://doi.org/10.1109/imccc.2018.00079
26. Abramova, V. V. (2015). A blind method for additive noise variance evaluation based on homogeneous region detection using the fourth central moment analysis. Telecommunications and Radio Engineering, 74 (18), 1651–1669. https://doi.org/10.1615/telecomradeng.v74.i18.50
27. Lukin, V., Kryvenko, S., Bondzulic, B., Bujakovic, N. (2024). Compression Ratio Behavior for BPG-based Compression of Grayscale Images. Proceedings of ATIT. Lviv. (in print).
28. Lukac, R., Plataniotis, K. N. (Eds.) (2007). Color Image Processing: Methods and Applications. Image processing series. Boca Raton: CRC/Taylor & Francis.
29. Jayachandran, S. (2017). Digital imaging in dentistry: A review. Contemporary Clinical Dentistry, 8 (2), 193–194. https://doi.org/10.4103/ccd.ccd_535_17
30. Krivenko, S., Lukin, V., Krylova, O., Kryvenko, L., Egiazarian, K. (2020). A Fast Method of Visually Lossless Compression of Dental Images. Applied Sciences, 11 (1), 135. https://doi.org/10.3390/app11010135
31. Flint, A. C. (2012). Determining optimal medical image compression: psychometric and image distortion analysis. BMC Medical Imaging, 12 (1). https://doi.org/10.1186/1471-2342-12-24
32. Lin, W., Jay Kuo, C.-C. (2011). Perceptual visual quality metrics: A survey. Journal of Visual Communication and Image Representation, 22 (4), 297–312. https://doi.org/10.1016/j.jvcir.2011.01.005
33. Chandler, D. M. (2013). Seven Challenges in Image Quality Assessment: Past, Present, and Future Research. ISRN Signal Processing, 2013, 1–53. https://doi.org/10.1155/2013/905685
34. Moorthy, A. K., Bovik, A. C. (2010). Visual quality assessment algorithms: what does the future hold? Multimedia Tools and Applications, 5 1(2), 675–696. https://doi.org/10.1007/s11042-010-0640-x
35. Andrew, B. W. (1993). DCTune: A technique for visual optimization of DCT quantization matrices for individual images. Society for Information Display Digest of Technical Papers, 946–949.
36. Bosse, S., Maniry, D., Muller, K.-R., Wiegand, T., Samek, W. (2018). Deep Neural Networks for No-Reference and Full-Reference Image Quality Assessment.

IEEE Transactions on Image Processing, 27 (1), 206–219. https://doi.org/10.1109/tip.2017.2760518
37. Lukin, V., Abramov, S., Krivenko, S., Kurekin, A., Pogrebnyak, O. (2013). Analysis of classification accuracy for pre-filtered multichannel remote sensing data. Expert Systems with Applications, 40 (16), 6400–6411. https://doi.org/10.1016/j.eswa.2013.05.061
38. Ponomarenko, N., Silvestri, F., Egiazarian, K., Carli, M., Astola, J., Lukin, V. (2007). On between-coefficient contrast masking of DCT basis functions. CD-ROM Proceedings of VPQM, 4.
39. Lukin, V., Ponomarenko, N., Egiazarian, K., Astola, J. (2015). Analysis of HVS-Metrics' Properties Using Color Image Database TID2013. Proceedings of ACIVS. Italy, 613–624. https://doi.org/10.1007/978-3-319-25903-1_53
40. Ieremeiev, O. I., Lukin, V. V., Ponomarenko, N. N., Egiazarian, K. O., Astola, J. (2016). Combined full-reference image visual quality metrics. Electronic Imaging, 28 (15), 1–10. https://doi.org/10.2352/issn.2470-1173.2016.15.ipas-180
41. Ponomarenko, N., Ieremeiev, O., Lukin, V., Egiazarian, K., Jin, L., Astola J. (2013). Color Image Database TID2013: Peculiarities and Preliminary Results, Proceedings of EUVIP. Paris, 106–111.
42. Varga, D. (2020). A Combined Full-Reference Image Quality Assessment Method Based on Convolutional Activation Maps. Algorithms, 13 (12), 313. https://doi.org/10.3390/a13120313
43. Ieremeiev, O., Lukin, V., Okarma, K., Egiazarian, K. (2020). Full-Reference Quality Metric Based on Neural Network to Assess the Visual Quality of Remote Sensing Images. Remote Sensing, 12 (15), 2349. https://doi.org/10.3390/rs12152349
44. Okarma, K.; Rutkowski, L., Scherer, R., Tadeusiewicz, R., Zadeh, L. A., Zurada, J. M. (Eds.) (2010). Combined Full-Reference Image Quality Metric Linearly Correlated with Subjective Assessment. Artificial Intelligence and Soft Computing. Berlin, Heidelberg: Springer, 539–546. https://doi.org/10.1007/978-3-642-13208-7_67
45. Li, F., Ieremeiev, O., Lukin, V., Egiazarian, K. (2024). BPG-Based Lossy Compression of Three-Channel Remote Sensing Images with Visual Quality Control. Remote Sensing, 16 (15), 2740. https://doi.org/10.3390/rs16152740
46. Ziaei Nafchi, H., Shahkolaei, A., Hedjam, R., Cheriet, M. (2016). Mean Deviation Similarity Index: Efficient and Reliable Full-Reference Image Quality Evaluator. IEEE Access, 4, 5579–5590. https://doi.org/10.1109/access.2016.2604042
47. Reisenhofer, R., Bosse, S., Kutyniok, G., Wiegand, T. (2018). A Haar wavelet-based perceptual similarity index for image quality assessment. Signal Processing: Image Communication, 61, 33–43. https://doi.org/10.1016/j.image.2017.11.001

48. Zhang, L., Zhang, L., Mou, X., Zhang, D. (2011). FSIM: A Feature Similarity Index for Image Quality Assessment. IEEE Transactions on Image Processing, 20 (8), 2378–2386. https://doi.org/10.1109/tip.2011.2109730
49. Ieremeiev, O., Lukin, V., Okarma, K., Egiazarian, K., Vozel, B. (2022). On properties of visual quality metrics in remote sensing applications. Electronic Imaging, 34 (10), 354-1-354–356. https://doi.org/10.2352/ei.2022.34.10.ipas-354
50. Barman, N., Martini, M. G. (2020). An Evaluation of the Next-Generation Image Coding Standard AVIF. 2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX). Athlone: IEEE. https://doi.org/10.1109/qomex48832.2020.9123131
51. Lainema, J., Hannuksela, M. M., Vadakital, V. K. M., Aksu, E. B. (2016). HEVC still image coding and high efficiency image file format. 2016 IEEE International Conference on Image Processing (ICIP), 71–75. https://doi.org/10.1109/icip.2016.7532321
52. Taubman, D. S., Marcellin, M. W. (2002). JPEG2000 Image Compression Fundamentals, Standards and Practice. Springer US. https://doi.org/10.1007/978-1-4615-0799-4
53. Beong-Jo Kim, Pearlman, W. A. (1997). An embedded wavelet video coder using three-dimensional set partitioning in hierarchical trees (SPIHT). Proceedings DCC'97. Data Compression Conference. Snowbird: IEEE Comput. Soc. Press, 251–260. https://doi.org/10.1109/dcc.1997.582048
54. Ponomarenko, N., Lukin, V., Egiazarian, K., Astola, J. (2005). DCT Based High Quality Image Compression. Image Analysis. Joensuu, 1177–1185. https://doi.org/10.1007/11499145_119
55. Ponomarenko, N. N., Egiazarian, K. O., Lukin, V. V., Astola, J. T. (2007). High-Quality DCT-Based Image Compression Using Partition Schemes. IEEE Signal Processing Letters, 14 (2), 105–108. https://doi.org/10.1109/lsp.2006.879861
56. Lukin, V., Abramov, S., Kozhemiakin, R., Rubel, A., Uss, M., Ponomarenko, N. et al.; Emre Celebi, M., Lecca, M., Smolka, B. (2015). DCT-Based Color Image Denoising: Efficiency Analysis and Prediction. Color Image and Video Enhancement. Springer, 55–80. https://doi.org/10.1007/978-3-319-09363-5_3
57. Zriakhov, M. S., Lukin, V. V. (2005). Obespechenie zadannogo kachestva pri szhatii izobrazhenii s poteriami. Radiotekhnika, 143, 76–82.
58. Zemliachenko, A. N., Kolganova, O. E., Lukin, V. V. (2011). Acceleration image compression with required visual quality. Radioelektronnye i kompiuternye sistemy, 4 (52), 52–59.

59. Kryvenko, L., Krylova, O., Lukin, V., Kryvenko, S. (2024). Intelligent visually lossless compression of dental images. Advanced Optical Technologies, 13. https://doi.org/10.3389/aot.2024.1306142
60. Li, F. (2023). Design and analysis of efficient methods for providing a desired quality in image lossy compression. [Extended abstract of PhD thesis; National Aerospace University].
61. Li, F., Krivenko, S., Lukin, V.; Nechyporuk, M., Pavlikov, V., Kritskiy, D. (Eds.) (2021). A Fast Method for Visual Quality Prediction and Providing in Image Lossy Compression by SPIHT. Integrated Computer Technologies in Mechanical Engineering – 2020. Cham: Springer International Publishing, 17–29. https://doi.org/10.1007/978-3-030-66717-7_2
62. Li, F., Krivenko, S., Lukin, V. (2020). Adaptive two-step procedure of providing desired visual quality of compressed image. Proceedings of the 2020 4th International Conference on Electronic Information Technology and Computer Engineering. Xiamen, 407–414. https://doi.org/10.1145/3443467.3443791
63. Li, F., Krivenko, S., Lukin, V. (2020). An Approach to Better Portable Graphics (BPG) Compression with Providing a Desired Quality. 2020 IEEE 2nd International Conference on Advanced Trends in Information Theory (ATIT), 13–17. https://doi.org/10.1109/atit50783.2020.9349289
64. Li, F., Lukin, V. V., Okarma, K., Fu, Y., Duan, J.; Chen, C.-H. (Ed.) (2022). Intelligent Lossy Compression Method of Providing a Desired Visual Quality for Images of Different Complexity. Applied Mathematics, Modeling and Computer Simulation. IOS Press. https://doi.org/10.3233/atde220050
65. Li, F., Abramov, S., Dohtiev, I., Lukin, V. (2024). Advantages and drawbacks of two-step approach to providing desired parameters in lossy image compression. Advanced Information Systems, 8 (1), 57–63. https://doi.org/10.20998/2522-9052.2024.1.07
66. Li, F., Krivenko, S., Lukin, V. (2020). A Two-step Approach to Providing a Desired Visual Quality in Image Lossy Compression. 2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET). Lviv-Slavske: IEEE, 502–506. https://doi.org/10.1109/tcset49122.2020.235483
67. Li, F., Krivenko, S., Lukin, V. (2020). Two-step providing of desired quality in lossy image compression by SPIHT. Radioelectronic and Computer Systems, 2, 22–32. https://doi.org/10.32620/reks.2020.2.02
68. Kozhemiakin, R., Lukin, V., Vozel, B. (2017). Image quality prediction for DCT-based compression. 2017 14th International Conference the Experience

of Designing and Application of CAD Systems in Microelectronics (CADSM). Lviv-Polyana: IEEE 225–228. https://doi.org/10.1109/cadsm.2017.7916121
69. Vozel, B., Kozhemiakin, R. A., Abramov, S. K., Lukin, V. V., Chehdi, K.; Bruzzone, L., Bovolo, F., Benediktsson, J. A. (Eds.) (2017). Output MSE and PSNR prediction in DCT-based lossy compression of remote sensing images. Image and Signal Processing for Remote Sensing XXIII. Warsaw: SPIE, 84. https://doi.org/10.1117/12.2278002
70. Krivenko, S., Zriakhov, M., Lukin, V., Vozel, B. (2018). MSE prediction in DCT-based lossy compression of noise-free and noisy remote sensing. 2018 14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET). Lviv-Slavske: IEEE, 883–888. https://doi.org/10.1109/tcset.2018.8336338
71. Minguillo'n, J., Pujol, J. (2001). JPEG standard uniform quantization error modeling with applications to sequential and progressive operation modes. Journal of Electronic Imaging, 10 (2), 475–485. https://doi.org/10.1117/1.1344592
72. Abramova, V., Lukin, V., Abramov, S., Kryvenko, S., Lech, P., Okarma, K. (2023). A Fast and Accurate Prediction of Distortions in DCT-Based Lossy Image Compression. Electronics, 12 (11), 2347. https://doi.org/10.3390/electronics12112347
73. Abramova, V., Lukin, V., Abramov, S., Abramov, K., Bataeva, E. (2022). Analysis of Statistical and Spatial Spectral Characteristics of Distortions in Lossy Image Compression. 2022 IEEE 2nd Ukrainian Microwave Week (UkrMW). Kharkiv, 644–649. https://doi.org/10.1109/ukrmw58013.2022.10036949
74. Li, F., Kryvenko, S., Lukin, V.; Urbach, H. P., Jiang, H. (Eds.) (2023). Remote Sensing Image Lossy Compression Based on JPEG with Controlled Visual Quality. Proceedings of the 7th International Symposium of Space Optical Instruments and Applications. Singapore: Springer, 8–19. https://doi.org/10.1007/978-981-99-4098-1_2
75. Bondžulić, B., Stojanović, N., Petrović, V., Pavlović, B., Miličević, Z. (2021). Efficient Prediction of the First Just Noticeable Difference Point for JPEG Compressed Images. Acta Polytechnica Hungarica, 18 (8), 201–220. https://doi.org/10.12700/aph.18.8.2021.8.11
76. Lin, H., Chen, G., Jenadeleh, M., Hosu, V., Reips, U.-D., Hamzaoui, R., Saupe, D. (2022). Large-Scale Crowdsourced Subjective Assessment of Picturewise Just Noticeable Difference. IEEE Transactions on Circuits and Systems for Video Technology, 32 (9), 5859–5873. https://doi.org/10.1109/tcsvt.2022.3163860
77. Li, F., Lukin, V., Kryvenko, S., Bondzulic, B., Bujakovic, D., Pavlovic, B. (2023). Strange Images in Remote Sensing and Their Properties. Ukrainian Journal of Remote Sensing, 10 (2), 12–18. https://doi.org/10.36023/ujrs.2023.10.2.240

78. Pavlović, B., Bondžulić, B., Stojanović, N., Petrović, V., Bujaković, D. (2023). Prediction of the First Just Noticeable Difference Point based on Simple Image Features. 2023 Zooming Innovation in Consumer Technologies Conference (ZINC). Novi Sad, 125–130. https://doi.org/10.1109/zinc58345.2023.10173865
79. Bondžulić, B., Stojanović, N., Lukin, V., Stankevich, S. A., Bujaković, D., Kryvenko, S. (2024). Target acquisition performance in the presence of JPEG image compression. Defence Technology, 33, 30–41. https://doi.org/10.1016/j.dt.2023.12.006
80. Bondžulić, B., Pavlović, B., Stojanović, N., Petrović, V. (2022). Picture-wise just noticeable difference prediction model for JPEG image quality assessment. Vojnotehnicki Glasnik, 70 (1), 62–86. https://doi.org/10.5937/vojtehg70-34739

CHAPTER 2

Solving pattern recognition problems in shipping monitoring systems based on CNN-models

Pavlo Mykhalichenko
Tetiana Cherniavska
Bohdan Cherniavskyi
Viktor Nadtochii
Anatolii Nadtochyi
Maiia Korkh

Abstract

The chapter examines modern approaches to solving pattern recognition problems in shipping monitoring systems using convolutional neural network (CNN) models. Traditional rule-based and classical classification methods remain ineffective for complex visual tasks in inland and maritime transport due to high variability of vessel shapes, viewing angles, illumination conditions, and significant intraclass differences. Deep learning architectures – particularly CNNs – are shown to provide robust and scalable solutions by automatically extracting hierarchical features and forming stable, noise-resistant image representations.

The chapter provides a comprehensive review of the fundamental principles of CNN construction, including convolutional layers, feature maps, activation functions, pooling operations, regularization techniques, batch normalization, and optimization strategies. Special attention is devoted to the mechanisms of training neural networks: gradient-based optimization, stochastic gradient descent, hyperparameter tuning, prevention of overfitting, and the use of computational graphs for backpropagation.

Experimental evaluation was conducted using a dataset comprising more than 12,000 annotated images of inland waterway vessels, including tugboats, barges, passenger ships, and dry cargo vessels, captured in real river-port conditions under varying illumination, viewpoints, partial occlusions, and noise. The proposed CNN model achieved an overall classification accuracy of 93.4%, with precision of 92.1%, recall of 91.7%, and an F1-score of 91.9%. Comparative analysis demonstrates that

 DOI: 10.21303/978-9908-845-05-0.ch2

the proposed approach outperforms classical methods such as k-nearest neighbors, linear classifiers, and feature-based support vector machines by 8–12% on average.

The practical importance of CNN-based methods for real-time monitoring of river port water areas in Ukraine is emphasized. The proposed models enable automated classification and recognition of surface vessels, enhancing the efficiency, accuracy, and adaptability of intelligent surveillance systems. The results contribute to the development of advanced decision-support tools in maritime transport, improving safety, environmental monitoring, and operational management of shipping flows.

Keywords

Convolutional neural networks, pattern recognition, vessel classification, shipping monitoring systems, deep learning, computer vision, feature extraction, optimization, batch normalization, river transport, visual information processing, machine learning, transport security, environmental monitoring.

2.1 Introduction

The novelty of this work lies not in proposing a fundamentally new convolutional neural network architecture, but in the targeted adaptation and systematic validation of CNN-based approaches for an underexplored application scenario – vision-based monitoring and classification of inland waterway vessels in river-port environments.

Unlike most existing studies that focus on open-sea maritime surveillance, satellite or SAR imagery, or multimodal fusion of AIS and visual data, this chapter addresses vessel type classification based solely on visible-spectrum images acquired from shore-based cameras operating under real port conditions.

The key novel aspects of the proposed approach include:

– its application to inland river ports characterized by limited camera elevation, variable viewpoints, dense object layouts, and complex backgrounds;

– its focus on vessel type classification rather than generic ship detection;

– a quantitative comparison with classical classifiers under identical inland waterway conditions;

– its practical orientation toward real-time port monitoring systems without reliance on AIS data.

A critical analysis of prior studies shows that while existing deep learning-based approaches demonstrate high performance in maritime and satellite scenarios, their direct applicability to inland waterways is limited due to differences in scene

geometry, data availability, and operational constraints. The proposed approach addresses these limitations, thereby extending the state of the art to a new and practically significant domain.

Within this chapter, the study is focused on the problem of visual classification of inland waterway vessels in port water areas based on images acquired from shore-based video surveillance systems operating in the visible spectrum. The primary attention is given to distinguishing the main types of river transport vessels, including tugboats, barges, passenger ships, and dry cargo vessels, under real operating conditions of port infrastructure. The study is deliberately limited to the analysis of two-dimensional static images without the use of temporal information and without incorporating auxiliary data sources such as automatic identification system (AIS) messages, radar measurements, or satellite imagery. This approach corresponds to the typical operating conditions of local port monitoring systems, where visual information from shore-based cameras is the primary available data source.

The chosen methodological approach is based on the use of convolutional neural networks, which is motivated by the limited effectiveness of classical computer vision methods in vessel recognition tasks under conditions of high intra-class variability, complex background environments, varying illumination, and constrained viewing angles. In contrast to approaches relying on handcrafted features and heuristic rules, convolutional neural networks enable automatic extraction of hierarchical visual feature representations that are robust to these factors. The focus on vessel type classification, rather than detection or tracking, is determined by the practical requirements of port monitoring systems, where the presence of an object has already been established and the key task is rapid identification to support operational decision-making. At the same time, it should be noted that the obtained results may require further adaptation when applied to other regions or in scenarios involving video streams or multimodal data sources.

2.2 Solving problems of classification of visual information

Computer vision (CV) is the analysis of visual data. The volume of this data in the world is constantly growing. About 80% of all Internet traffic is video – and this is without taking into account images and other types of visual information. Therefore, it is important to develop algorithms that can understand and process this data.

Visual information is sometimes compared to dark matter by analogy with physics. Dark matter makes up a very large fraction of the mass of the Universe, but it is not directly possible to observe it. Visual data is about the same: it contains a lot

of bits flying around the Internet. But it is very difficult for CV algorithms to understand what they actually consist of.

Visual perception affects many fields of science and technology: in physics it is important to understand the process of image formation, in biology and psychology scientists' study how people and animals perceive and process visual information. In robotics and automotive engineering, images help with orientation in the territory. To create systems that implement visual perception algorithms, knowledge in the fields of computer science, mathematics and design is required.

The main task that CV focuses on is image classification. That is, the algorithm must assign an image to one of the previously known categories. This problem is being solved both in research circles and in industry [1].

The engine of progress in solving this problem was the recently invented convolutional neural networks (CNNs). The real breakthrough came in 2012, when G. Hinton and his graduate students A. Kryzhevsky and I. Sutzkever created a seven-layer ANN AlexNet. It performed very well in the ImageNet competition. Since then, there has been a trend towards creating ever deeper networks. In 2014, the multilayer GoogLeNet and VGG appeared. At the same time, a paper was published by researchers from Microsoft Research Asia, who created a residual neural network (RN) with 152 layers.

The main breakthrough in the development of neural architectures occurred only a few years ago. This is primarily due to the emergence and rapid growth in the performance of graphics processors GPU [2]. These processors allow processing huge amounts of data in parallel, which makes them an ideal tool for large-scale computing and training RN. In addition, the amount of data available for training has increased significantly.

The CV opens up a lot of potential tasks, the main one of which is to teach a machine to see the world like a human. This goal is still unattainable, as it gives rise to a lot of related problems: for example, recognizing human activity from video, accurate 3D reconstruction of objects, semantic image segmentation, and many others. But as it is constantly moved forward and invent such amazing things as augmented and virtual reality, let's probably come up with new, interesting solutions [1].

The current state of research in vessel recognition for maritime and inland waterway monitoring is characterized by a rapid transition from traditional computer vision algorithms to deep learning-based approaches. Several recent surveys confirm that convolutional neural networks and their derivatives have become the de facto standard for ship detection and classification in satellite imagery, aerial data, and video streams. M. J. Er and Y. Zhang, for example, provide a comprehensive review of deep learning-based ship detection techniques, ranging from early CNN models

to modern one-stage detectors such as the YOLO and SSD families, and consistently demonstrate that deep architectures outperform classical methods in terms of accuracy and robustness to noise and background clutter [3].

Further progress has been driven by dedicated models for ship detection and tracking in video. B. Zhang et al. presents an extended survey of deep-learning methods for automatic ship detection and tracking, highlighting four significant challenges: environmental variability, multi-scale targets, occlusions, and the need for lightweight models capable of real-time performance [4]. For synthetic aperture radar (SAR) imagery, several surveys emphasize the specific difficulties related to speckle noise, background complexity, and limited spatial resolution, and point to an increasing adoption of advanced deep models, including transformer-based architectures [5].

In parallel, a new class of CNN-Transformer hybrid models have emerged. D. Liu [6] introduces transformer-oriented detectors (e.g. TS2Anet and DETR-like frameworks) tailored to ship detection in complex maritime environments, such as congested ports and high-clutter coastal regions [6]. More recently, Y. Wang and X. Li proposed Ship-DETR, a transformer-based detector designed for efficient ship detection in challenging maritime scenes, demonstrating the advantages of attention mechanisms in handling irregular spatial structures and densely populated scenes [7].

Another important line of work focuses on multimodal fusion, which combines visual data (such as RGB or SAR imagery) with Automatic Identification System (AIS) messages and other sensor sources. Studies by A. Galdelli et al. demonstrate that integrating heterogeneous data streams can significantly enhance the reliability of ship detection and classification, notably when individual channels are partially missing or severely degraded [8]. However, most of these approaches are tailored explicitly to open-sea scenarios and satellite or airborne platforms, where the observation scale and typical traffic patterns differ substantially from those in inland waterways and river ports.

In contrast to the majority of existing research that focuses on maritime domains, a more petite but growing body of work addresses inland waterways. M. Salem et al. proposes a CNN-based method for classifying vessel types on inland waterways. Yet, their experiments rely on relatively standardized imagery and do not fully reflect the complexity of port environments [9]. Studies by G. Agorku emphasize the significance of deep learning for barge load estimation, autonomous vessel control, and performance assessment of inland waterway freight corridors, primarily in the context of AIS-based trajectory analysis rather than vision-based recognition [10].

Within this landscape, the approach presented in this chapter focuses specifically on **vision-based classification of inland waterway vessels in real river-port**

conditions in Ukraine, using images captured by shore-based cameras operating in the visible spectrum. Unlike most prior work, which targets satellite surveillance or generic maritime scenes, the proposed CNN model is explicitly optimized for:

– varying viewpoints typical of shore-side infrastructure;
– complex and cluttered backgrounds (piers, port facilities, other vessels);
– limited sensor height and constrained fields of view;
– the characteristic mix of inland and mixed navigation vessel types.

Consequently, this chapter advances the state of the art in two key directions:

1. It demonstrates the practical applicability of CNN-based classification to **river-port monitoring on inland waterways**, a scenario that remains relatively underexplored in literature.

2. It provides a **quantitative evaluation** on a real dataset collected from Ukrainian river ports, enabling a direct comparison between CNN-based methods and classical classifiers discussed earlier in the chapter.

The experimental study described in this chapter is based on a combination of publicly available datasets and a proprietary image collection obtained under real river-port operating conditions. For preliminary testing and architectural validation, well-known open benchmark datasets for image classification were used, including CIFAR-10, which provides standardized RGB images and is commonly applied for verifying the learning capability of convolutional neural networks. The use of such benchmark datasets ensures methodological consistency with prior studies and enables reproducibility of baseline experiments.

The primary dataset for vessel classification was formed from images acquired by shore-based video surveillance cameras installed in river-port areas. This dataset includes visual samples of inland waterway vessels such as tugboats, barges, passenger ships, and dry cargo vessels captured under varying illumination conditions, viewing angles, background clutter, and partial occlusions. Due to operational and security constraints, this dataset is not publicly available; however, its structure, class composition, and acquisition conditions are described in sufficient detail to enable replication using similar port monitoring systems.

All computational experiments were conducted using widely adopted open-source software tools. The implementation of neural network models was performed using the Python programming language and deep learning frameworks such as PyTorch, which provide transparent model definitions, reproducible training procedures, and extensive community support. Image preprocessing and augmentation operations were carried out using standard computer vision libraries, including OpenCV and NumPy. The use of open-source tools ensures transparency of the experimental pipeline and facilitates independent reproduction of the proposed approach.

When the system receives an input image, it already knows a fixed set of categories or labels. These can be any objects: "tugboats", "barges", "passenger ships", "dry cargo ship". The computer must look at the image and assign it one of the labels.

From the outside, the task seems simple, since most of our visual system is programmed to recognize objects. But for a machine, it is not so simple. Especially in such a specific field as water transport, since the outlines of vessels are similar and it is necessary to determine certain classification features of each type of vessel.

When a computer analyzes an image, it doesn't "see" the whole picture of a cat or, for example, a tugboat. It only "sees" a giant grid of numbers. For example, if the image size is 800 by 600 and each pixel is represented by three numbers for the red, green, and blue channels, the resulting grid is $800 \times 600 \times 3 = 1{,}440{,}000$ numbers. It is very difficult to distinguish any particular object represented in the image from this grid.

This problem is called the "semantic gap" – a misunderstanding of the information contained in the data. For example, if to photograph a dry cargo ship from a different angle or in different lighting, the entire grid of numbers will look completely different. In addition, the photograph may only show part of the ship, for example the stern. Recognition algorithms must be resistant to such changes.

In addition to these difficulties, there is the problem of intraclass variation, where a single concept encompasses a multitude of visual manifestations. For example, passenger ships can be of different types, ages, and sizes. And recognition methods must handle all possible variations.

The first thing that comes to mind is to create reference rules. It is known that ships have certain deck equipment. In photos of them, it is possible to detect edges and then classify different angles and boundaries: for example, determine how the lines of the hull and bow of a ship meet.

But explicit rule sets don't work very well: any deviation can break everything, and new objects will have to be created with new conditions. Therefore, this approach is not scalable.

Instead of trying to manually create a set of rules, it is possible to open the Internet and collect a large dataset with photos of different types of river and sea vessels. For this, Google image search or a ready-made dataset is suitable. Then it is necessary to train the classifier by sending all the collected images to it. At the output, let's get a model that generalizes the knowledge of recognizing different objects. After that, it can work on new images and distinguish tugboats from dry docks.

So, instead of one function that simply recognizes an object in an input image, there are two: the first is called "training" – this is the process of processing images and creating a model. The second function – "prediction" – recognizes new photos. Together, they form the basis for CNN and DNN in general.

To begin, let's look at the simplest classifier, which is called the "nearest neighbor method". During the training process, it remembers all the original data, and during the prediction stage, it tries to find the most similar new images.

There are many different ways to compare two images. In the example below, let's use the $L_{1\ distance}$, also known as the Manhattan distance. It simply compares the pixels of the images. Let's say that there is a test sample of 4 × 4 pixels. Let's take one of the training images and calculate the absolute difference between the colors of the pixels of the training and test samples, the $L_{1\ distance}$

$$d_1(I_1, I_2) = \sum_p \left| I_1^p - I_2^p \right|, \tag{2.1}$$

and then sum up the obtained values.

If the training set contains N examples, then training will always be performed in constant time $p(1)$ [12], and prediction will be performed in linear time $p(N)$ because the test image is compared to each training example. Linear prediction time is not very good. In reality, classifiers should be trained slowly and tested quickly. More advanced algorithms work in this way: they can be trained for a long time in a data center or on cloud servers, and then run on mobile phones.

The l_2 *(Euclidean)* distance is also often used in classification problems. It looks like the square root of the sum of the squared differences between pixel colors.

Distance metrics make different assumptions about the expected geometry or topology of the space. l_1 forms a square region, while l_2 creates a circle. When the coordinate system is rotated, the Euclidean distance l_2 will not change, while the Manhattan (MN) l_1 will give a different result. It is important to take this spatial effect into account and choose the metric according to the original task.

The process of choosing values that affect the performance of any method, such as the number of neighbors k and the distance metric, is called hyperparameter tuning. Hyperparameters cannot be explicitly extracted from the training data and depend only on the algorithm itself, so there are no clear recommendations for their selection. Most often, it is possible to find values by trial and error, finding out which ones work best.

Let's figure out what "best" is:

1. Choose parameters that will give the highest accuracy on the training data. And this is a very bad idea. In the case of the nearest neighbor method, at $k = 1$, almost perfect accuracy is achieved during training, but the algorithm does extremely poorly on the test data. This is called "overtraining".

2. Split the population into training and testing sets and find hyperparameters that will make the algorithm perform better on the test samples. This strategy looks smarter, but in reality, it is also very bad. The main idea of L_1 MN is that it is possible

to know how the method will perform. And if to choose parameters that achieve good results on known images, then there is no guarantee that they will be achieved on other unknown images.

3. Split the dataset into training, testing, and validation sets; select hyperparameters for the evaluation data and test them on the test data. And that's a good idea. First, the classifier is trained with different parameter options. Then, the values that work best on the evaluation data are selected. After that, the model processes the test set only once. The accuracy achieved in this case shows the true efficiency of the classifier.

4. Divide the training data into many small subsamples, then cross-validate using different subsamples as the evaluation data, and average the results.

The method is called "cross-validation" and works well on small datasets, but requires a lot of computational resources for huge datasets.

The *k*-nearest neighbors' method is never applied to photographs because: it is very slow on test data; the metric of distances between pixel colors does not indicate the similarity of images.

Another problem is the so-called curse of dimensionality, which is associated with the exponential growth of the amount of data as the dimensionality of the space increases. Therefore, algorithms based on brute force become inefficient as the dimensionality of the system increases.

Linear classification algorithms are quite simple. However, they are used to create full-fledged neural networks. It is similar to Lego: it is possible to put different components together and build a "tower" convolutional neural network (CNN).

Let's use the CIFAR-10 dataset. To classify an image, it is possible to create a simple parametric model with two components. The first is the input data, usually denoted as *X*, and the second is a set of parameters or weights *W*. Now let's write a function that takes in the data *X* and the parameters *W* (**Fig. 2.1**) and then outputs 10 numbers describing the scores for each of the 10 categories in CIFAR-10.

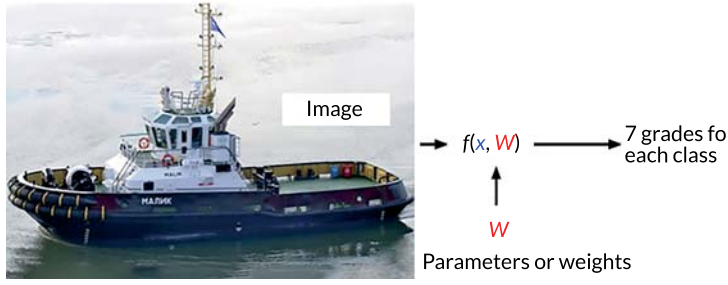


Fig. 2.1 Creating a linear estimate

A properly working model in the example above will give the highest score for the cat class. In this parametric approach, unlike the nearest neighbor method, let's generalize knowledge about the training data and use it to create parameters. Therefore, during testing, it was not possible to refer to the training sample every time.

To get a function that combines weights and data, the easiest way is to simply multiply them. This will be a linear classifier

$$f(x, W) = Wx + b, \tag{2.2}$$

where b – a linear vector, whose size is equal to the number of classes (in this case 10). It does not interact with the training data and produces independent predictions for the classes. For example, if the set contains many more images of tugboats than of dry cargo ships, then the elements of the bias vector corresponding to the cat class will be higher than the others.

The score for any class is the product of the image pixels and the corresponding rows of the weight matrix, with an offset added. This is similar to a pattern match: each row corresponds to some template image, and the score indicates its similarity to the original photo.

If to try to fold the rows of the weight matrix back into the image, it is actually possible to get patterns collected from the data. A linear classifier uses only one pattern for each class. By creating more complex functions that relate the data and parameters, it is possible to learn more patterns and achieve better accuracy.

In the decision domain, each image is represented as a point in a multidimensional space, which the linear classifier tries to fit within the boundaries of a linear solution (**Fig. 2.2**). In other words, it will separate the categories from each other by straight lines.

Another challenge is the phenomenon of multimodality, which occurs in unevenly distributed data. For example, photographs of ships may show them from different sides [12]. In this case, isolated islands with ships facing right, left, and other directions appear in the decision-making area.

Despite the challenges, linear classification is a simple algorithm that is easy to interpret and implement.

The main result that the loss function gives is an estimate of how well the classifier works on the sample. But it is possible to remember that the main goal of MN is to make the algorithm work correctly on the test sample. Therefore, a method that copes well with the training data may not work at all on new objects. This is called "overtraining". Suppose that our "dataset" consists of some points. If to force the algorithm to adapt perfectly to each point with zero losses, then the classification

graph will turn into a winding curve (**Fig. 2.3**, curve 1). But this is a bad result, because it is about accuracy on the test sample, not on the training one. If to check the work on the test data (marked by the green squares in **Fig. 2.3**), then the blue curve will become completely wrong. Most likely, the classifier should find some curve (**Fig. 2.3**, curve 2) that would approximately correspond to both those and other data.

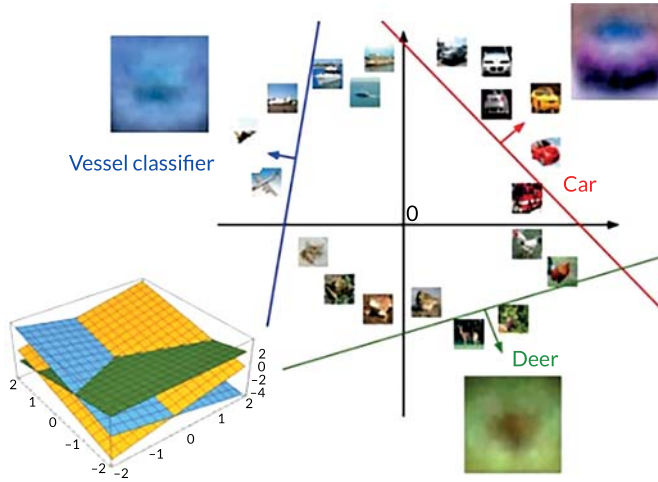


Fig. 2.2 Linear classification of a dataset

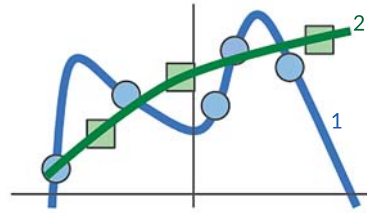


Fig. 2.3 Adding a test sample to the retrained model

This is the fundamental problem of MN. Various regularization methods are usually used to solve it.

It was said above that there are many options for optimal weights *W*. A regularization parameter is introduced specifically for this – it forces the model to choose such weights *W*, at which the solution will be the simplest. The concept of simplicity depends on the problem that the algorithm solves.

This idea is related to the principle of Occam's razor – if an observation can be explained by several hypotheses, the simplest one should be chosen. In the case of classification, the simplest partition of the data should be chosen – in the example above, this should be curve 2, not curve 1 of order *N*. So, our function now consists of two components: the training data loss (the sum of all L_i) and the regularization loss $R(W)$

$$L(W)=\frac{1}{N}\sum_{i=1}^{N}L_i\left(f\left(x_i,W\right),y_i\right)+\lambda R(W), \tag{2.3}$$

where λ – a hyperparameter that determines how strongly regularization will affect the model.

Now let's see what $R(W)$ functions are and how they work. In practice, quite a few types of regularization are used, but for the purposes of this chapter, let's consider only the most popular ones:

1. Regularization L_2.

It is used quite often in MN. It is the normal Euclidean norm of the parameters *W* (sometimes the square or half-square of the norm is taken). The idea is to simply add a "penalty" to the weights so that their values do not turn out to be too large. The function $R(W)$ looks like this

$$R(W)=\sum_k\sum_l W_{k,l}^2. \tag{2.4}$$

2. Regularization L_1.

L_1 uses the Manhattan norm *W* and also adds a penalty to the weights. With regularization the L_1 parameter matrix looks sparser

$$R(W)=\sum_k\sum_l\left|W_{k,l}\right|. \tag{2.5}$$

3. Elastic network.

This is a combination of regularization L_1 and L_2

$$R(W)=\sum_k\sum_l\beta W_{k,l}^2+\left|W_{k,l}\right|. \tag{2.6}$$

There are other types of regularizations, such as maximum norm, dropout, batch normalization, and stochastic depth. They are used in narrower GN problems, which will be explored in the following sections.

Let's imagine that we are walking through a large valley among mountains, fields and rivers. The height of each object in this landscape corresponds to the number of losses that appear at a certain setting *W*. Since the task is to achieve the

lowest losses, it is necessary to somehow find the lowest point in the valley. In practice, the optimization process looks like a set of iterative operations, where a random solution is taken as the starting point, which, using certain mathematical tools, is gradually improved.

Strategy 1. The first solution is to generate random values of *W* and see how well they work. But this is a bad idea and should not be used in practice. For example, for 10 classes of the CIFAR-10 "dataset", the probability of finding the corresponding *W* parameters will be 10%. This is a very bad indicator.

Strategy 2. Standing on a slope, it is possible to look for a descent. In mathematics, the analogue of finding a descent is the derivative. By taking a one-dimensional function *f*(*x*) and calculating its derivative at any point, it is possible to determine whether the function is increasing or decreasing at that point. But in MN, *X* is usually a vector, so the generalized analogue of descent is the vector of partial derivatives or gradient. The gradient indicates the direction of increase of a function, to find its decrease, a negative gradient is used.

So, now there is a parameter vector *W*, and our goal is to calculate the gradient vector ∇W. It is simply possible to take the derivatives of each individual element. This way finds out how much the losses will change if to move by an infinitesimal amount in one of the coordinate directions. This approach is called "numerical gradient". In most real cases, these calculations will be too slow. After all, the training data sometimes contains millions of examples, and the algorithms for processing them can be much more complex. To solve the problem, there is the so-called "analytic gradient". From the course of mathematical analysis, it is simple to write down the expression for the losses, and then use differential calculus and immediately find the necessary gradient. This becomes possible thanks to the analytical derivative, which does not require substitution of values, but works with the entire function at once [13]

$$\hat{f}(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = \lim_{\Delta x \to 0} \frac{\Delta f(x)}{\Delta x}. \quad (2.7)$$

It's actually better to use the analytical gradient, but check the solution numerically. This is called gradient checking.

First, let's initialize *W* with random values, calculate the losses and gradient, and then update the weights according to the negative direction of the gradient (**Fig. 2.4**).

"Step size" is a hyperparameter that specifies the step size moved with each new gradient calculation. It is sometimes called the Learning Rate (LR). This is one of the most important settings in MN.

```
while True:
  weights_grad = evaluate_gradient(loss_fun, data, weights)
  weights += - step_size * weights_grad
```

Fig. 2.4 Listing of implemented gradient descent in Python

Let's see how it looks visually. In the graph below, the large multi-colored area is our loss function (**Fig. 2.5**). The red area is the minimum values that is possible to achieve, and the blue and green indicate high losses.

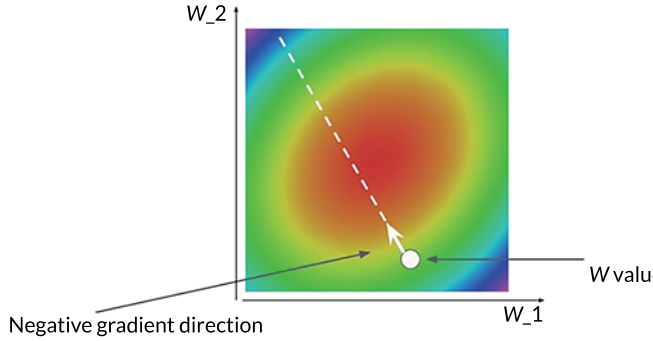


Fig. 2.5 Physical understanding of gradient descent

Let's start with the initial W at a random point and calculate the negative gradient direction that should lead step by step to the minimum loss. This is the most obvious example of stepwise gradient descent.

In the case of calculating the gradient W at each step, it is possible to consider all the training examples in the dataset. Because ∇W is the sum of the individual gradients caused by each training element. But their number can reach a million, and in this case, ordinary gradient descent will work very slowly. Therefore, stochastic gradient descent is often used in practice:

$$L(W)=\frac{1}{N}\sum_{i=1}^{N}L_i(x_i,y_i,W)+\lambda R(W); \tag{2.8}$$

$$\nabla_W L(W)=\frac{1}{N}\sum_{i=1}^{N}\nabla_W L_i(x_i,y_i,W)+\lambda\nabla_W R(W). \tag{2.9}$$

Its difference is that at each step the losses and gradient are calculated not on the entire "dataset", but on a small set of examples – a mini-package (and sometimes

only on one example). That is, the *W* parameters are updated after processing several objects, and not after passing through the entire dataset. This significantly speeds up the optimization process, and only one line is added to the code (**Fig. 2.6**):

```
while True:
  data_batch = sample_training_data(data, 256)
  weights_grad = evaluate_gradient(loss_fun, data_batch, weights)
  weights += - step_size * weights_grad
```

Fig. 2.6 Listing implementation of stochastic gradient descent in Python

The function *f* takes data *x* and parameter *W*, as input and outputs a vector of scores *s* for each of the categories to be classified. The loss function *L* (for example, SVM) determines how correct the scores are, with its help it is possible to calculate the data loss. Regularization helps to learn about the simplicity of the model. The goal is to find the parameters *W* that correspond to the smallest losses. Let's use the negative direction of the gradient of the function *L*, which indicates the path to its minimum. There are two main ways to calculate the gradient: analytical and numerical. The numerical method is quite simple, but works very slowly and gives approximate values [14]. The analytical gradient is more accurate and faster, but it is easy to make mistakes when calculating it. To avoid this and easily calculate the gradient even for complex functions, it is better to use computational graphs.

2.3 Development of a neural network for classifying visual information

In order to ensure the possibility of independent reproduction of the obtained results, this chapter provides a detailed description of the main stages involved in the construction and training of the convolutional neural network model for vessel classification. The overall solution pipeline follows a sequential structure that includes input image preprocessing, formation of training and test datasets, definition of the neural network architecture, model training using gradient-based optimization methods, and quantitative evaluation of classification performance.

During image preprocessing stage, input images are resized to a fixed spatial resolution, pixel intensity values are normalized, and basic data augmentation techniques are applied, when necessary, in order to improve robustness to illumination changes and viewpoint variations. The convolutional neural network architecture is designed according to a classical layered scheme that alternates convolutional

layers, nonlinear activation functions, and dimensionality reduction layers, followed by a fully connected classification block. Model training is performed using the back-propagation algorithm in combination with stochastic gradient descent and its commonly used variants.

The algorithmic sequence of model training and inference can be represented in the form of generalized pseudocode, reflecting the key computational steps of the process, including network parameter initialization, forward propagation, loss function computation, weight updates, and accuracy evaluation on a validation dataset. Such a representation enables the proposed approach to be reproduced using any modern deep learning framework.

The structural diagram of the convolutional neural network presented in this chapter illustrates the general logic of model construction and the interactions between its main components, serving as a reference template for implementing similar classification systems in water area monitoring applications. Taken together, the descriptions of the network architecture, training procedure, and experimental setup provide a sufficient level of detail to allow independent verification and replication of the reported results.

A computational graph is an illustrated representation of any function consisting of vertices (sometimes called nodes) and edges. Vertices are computational operations to be performed, and edges connect them in a certain sequence. **Fig. 2.7** shows an example of a graph with a classifier.

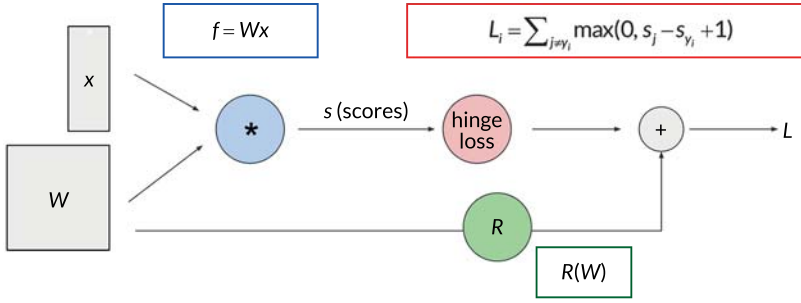


Fig. 2.7 Artificial neuron as a graph

A node with an operation (*) means the multiplication of the parameter matrices *W* and the data *x*, which results in a weight vector *s*. The next dependent loss node (hinge loss) defines the data loss *L*. Node *R* computes the regularization. Finally, it is possible to obtain the total loss by summing the regularization and data loss.

The advantage of graphs is that they allow to use the so-called backpropagation method. This algorithm recursively uses the differentiation rule of a complex function to calculate the gradient of each variable in the graph [15]. The method becomes very useful for really complex functions, such as those used in convolutional neural networks (CNNs).

Let's recall the linear classifier function. If to "transform" it on a neural network (NN), it is necessary to split the parameters W into two parts: W_1 and W_2 and apply one linear transformation on top of the other (**Fig. 2.8**).

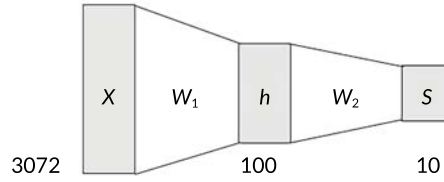


Fig. 2.8 Simple two-layer neural network

A simple two-layer NM with two linear layers was obtained

$$f = W_2 \max(0, W_1 x).$$

In **Fig. 2.8** above, x – the input data, h – the intermediate nonlinearity, and s – the output vector of estimates. In a broader sense, neural networks are complex functions made up of simple ones.

It was previously noted that each row of the weight matrix W is a template of one of the classes. These templates looked like an average object. It was also found that there is a problem with such single templates: for example, the class of ships in the datasets can be colored in different colors. Multilayer networks solve this problem: W_1 contains the single templates themselves, but now the estimates for them are stored in an intermediate nonlinear variable h. The next layer W_2 will combine the templates using a weighted sum, which will allow more accurate estimates of cars of other colors and other various objects.

By the way, nothing prevents from adding another layer to improve recognition accuracy

$$f = W_3 \max(0, W_2 \max(0, W_1 x)). \tag{2.10}$$

This is how deep NMs appear.

Convolutional neural networks (CNNs) handle large data sets well and are trained efficiently on GPUs using parallel computing. These features have become crucial to the fact that AI is now used almost everywhere. It solves problems such as image classification and search, object detection, segmentation, and is also used in more specialized fields of science and technology.

Let's assume that there is a 32 × 32 × 3 original 3D image. Let's stretch it into one long vector 3072 × 1 and multiply it with a weight matrix of size, for example, 10 × 3072. As a result, it is necessary to get an activation (an output with class scores) – to do this, let's take each of the 10 rows of the matrix and perform a scalar product with the original vector (**Fig. 2.9**).

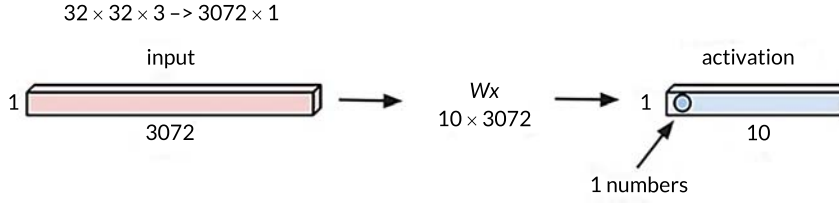


Fig. 2.9 Unfolding an image into a one-dimensional vector

The number 1 is the result of the dot product between one row of the weights *W* and the output vector. As a result, let's get a number that can be compared to the value of a neuron. In our case, let's get ten values. Fully connected layers work on this principle.

The main difference between convolutional layers is that they preserve the spatial structure of the image. Now it is possible to use weights in the form of small filters – spatial matrices that pass through the entire image and perform a dot product on each of its sections **Fig. 2.10**. In this case, the size of the filter always corresponds to the dimensions of the original image.

As a result of the image pass, let's obtain an activation map, also known as a feature map. This process is called spatial convolution [16, 17].

A large number of filters can be applied to an image and different activation maps can be obtained at the output. This way forms one convolutional layer. To create a whole NM, the layers are alternated one after the other, and activation functions (for example, ReLU [17]) and special pooling layers are added between them, which reduce the size of the feature maps.

In the first layers, convolutional filters are usually associated with low-level image features, such as edges and boundaries. In the middle, there are more complex

features, such as corners and circles. And in the final layers, the filters are more like some specific features that can be interpreted more broadly.

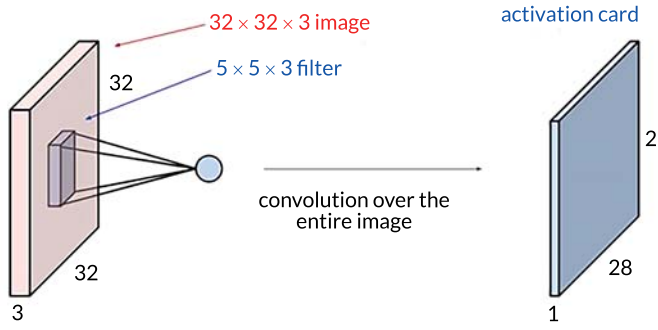


Fig. 2.10 Collapse of the image

Fig. 2.11 below shows examples of 5 × 5 filters and the resulting activation maps when applied to the original image (top left). The first filter (circled in red) looks like a small section of a border tilted to the right. If to apply it to a photograph, the highest values (white) will be where there are edges with roughly the same orientation. It is possible to see this by looking at the first activation map.

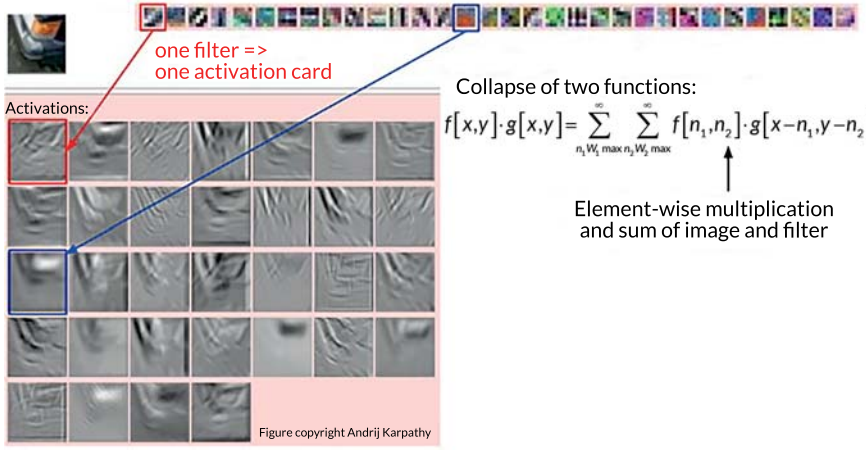


Fig. 2.11 Example of 5 × 5 filter operation

Thus, one layer of the NM finds the areas of the image that are most similar to the given filters. This process is very similar to the usual convolution of two functions. It shows how objects correlate with each other. Putting everything together, let's approximately get the following picture: taking the original photo, let's pass it through alternating convolutional layers, activation functions and pooling layers. At the end, let's use the usual fully connected layer, connected to all the outputs, which shows the final estimates of each class (**Fig. 2.12**).

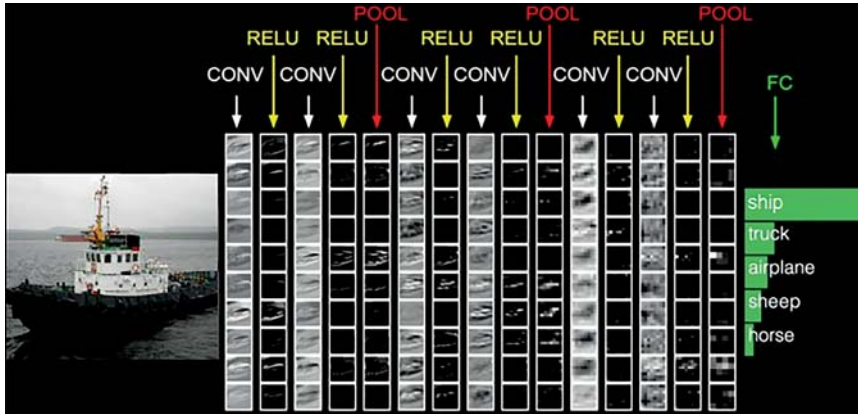


Fig. 2.12 Scheme of the convolutional operation neural networks

This is the principle that modern ZNMs work on. So, let's look at the full set of functions of the convolutional layer:

1. Accepts an original image of a certain dimension $W_1 \times H_1 \times D_1$.

2. Uses four hyperparameters for filters: the number of filters K; their size F; step S; number of extra zeros P (used to fill in the "lost" areas of the image after applying convolution).

3. Outputs an activation map of size $W_2 \times H_2 \times D_2$, where:

$$W_2 = (W_1 - F + 2P) / S + 1; \quad (2.11)$$

$$H_2 = (H_1 - F + 2P) / S + 1; \quad (2.12)$$

$$D_2 = K. \quad (2.13)$$

4. Uses $F \times F \times D_1$ weights per filter, for a total of $(F \times F \times D_1) \times K$ weights and K offsets.

Learning NM is an unpredictable and exciting process, which, however, requires careful preparation. In general, it can be divided into three main stages:

– **one-time setup**. Activation function selection, data preprocessing, weight initialization, regularization, gradient testing;

– **learning dynamics**. Tracking the learning process, optimizing and updating hyperparameters;

– **evaluation**. Using ensemble methods [18].

Activation function selection. Earlier it was found out that each layer of the NM receives input data. They are multiplied by the weights of the fully connected or convolutional layer, and the result is passed to the activation function or nonlinearity. Let's also talk about sigmoid and ReLU, which are often used as such functions. But the list of possible options is not limited to them. The task is to choose the activation function.

Data preparation. There are three most common methods of data preprocessing. Let's assume that the data X is a matrix of size $[N \times D]$:

1. *Subtracting the mean*. To avoid skewing the data and make it symmetric about zero, the mean value is subtracted from each element. This helps prevent the situation where all the original numbers turn out to be only positive or negative. In *NumPy*, the operation has the form $X = np.mean(X, axis = 0)$. In particular, when processing images, it is possible to subtract one value from all pixels (for example, $X = np.mean(X)$) or do it separately for each of the three-color channels.

2. *Normalization*. Transforming the data so that they are approximately the same scale. One option is to divide each dimension by its standard deviation: $(X = np.std(X, axis = 0))$. Another option is to normalize each value so that min and max are −1 and 1, respectively. Normalization should only be used if the original data has different formats or units. Pixel values do not fall outside the range of 0 to 255, so there is no need to perform normalization for them.

Initializing the weights. So, the neural network architecture was built and the data were prepared. Before starting training, it is necessary to initialize the parameters (weights).

What not to do: set the weights to zero. This will cause all neurons to behave the same way – not what we want. The NM should learn different features.

Small random variables. The most convenient option is to assign small values to the weights. Then all neurons will be unique and will gradually integrate into different parts of the network during the learning process. The implementation can look like this: $W = 0.01 \times np.random.randn(D, H)$. The $randn(n)$ method forms an array of size $n \times n$, the elements of which are random variables distributed according to the normal law with a mathematical expectation of 0 and a standard deviation

of 1 (Gaussian distribution [19]). The disadvantage of this is that it works well for small architectures, but copes much worse with bulky NMs.

Calibration using 1 / *sqrt*(*n*). The problem with the above method is that the variance of the random variables increases with the number of neurons. To avoid this, it is possible to scale the weights by dividing them by the square root of the number of inputs: *W* = *np.random.randn*(*n*) / *sqrt* (*n*). This ensures that all neurons in the network initially have approximately the same output distribution.

It is also possible to use the option *W* = *np.random.randn*(*n*) × *sqrt*(2.0 / *n*), which was proposed in one of the studies [20, 21]. It leads to the most successful distribution of neurons, so in practice it is possible to recommend using it.

Batch normalization. A method also known as batch normalization solves many initialization problems by forcing all activations to adopt a unit Gaussian distribution at the beginning of training.

Let's consider a small number of neuron activations on a layer. Let the activation function be a vector of dimension *D*: *x* = (*x*(1), …, *x*(*D*)). Let's normalize it by each dimension

$$\hat{x}^{(k)} = \frac{x^{(k)} - M\left(x^{(k)}\right)}{\sqrt{D\left(x^{(k)}\right)}}, \tag{2.14}$$

where *M*(*x*) – the mathematical expectation; *D*(*x*) – the variance, which is calculated over the entire training sample.

So, instead of initializing the weights, it is possible to use this simple differentiable function and get a normal distribution at each layer.

Batch normalization is usually applied between layers (fully connected or convolutional) and activation functions.

This is a very useful algorithm that is often used in modern MN. NMs that use batch normalization, are much more resistant to bad initialization.

Hyperparameter optimization. As is seen, training neural networks involves many stages of hyperparameter tuning. The most common are:

– initial SHN;

– the attenuation graph of the SH (for example, constant attenuation);

– regularization power.

If desired, it is even possible to modernize the network architecture if to suspect that it was not chosen very well.

The graphical representations of the experimental results are based on the quantitative performance indicators summarized in **Table 2.1**. As shown by the obtained

data, the proposed convolutional neural network–based model demonstrates a clear advantage over classical classification methods across all major evaluation metrics. In particular, the overall classification accuracy of the CNN approach reaches 93.4%, exceeding the performance of *k*-nearest neighbors, linear classifiers, and support vector machines with handcrafted features by approximately 7–12%.

Table 2.1 Quantitative performance metrics for vessel classification

Classification method	Overall accuracy, %	Precision, %	Recall, %	F1-score, %
k-nearest neighbors	81.2	78.5	76.9	77.7
Linear classification	84.6	83.1	82.4	82.7
SVM (handcrafted features)	86.1	84.8	83.9	84.3
CNN (proposed approach)	93.4	92.1	91.7	91.9

The precision and recall values indicate that the proposed model is capable of reliably distinguishing different vessel types under conditions of complex background scenes, varying illumination, and changing viewing angles that are typical for inland river-port environments. The aggregated F1-score confirms the balanced performance of the classifier and indicates the absence of a significant trade-off between false-positive and false-negative errors.

Thus, the presented numerical results directly substantiate the trends illustrated in the corresponding graphs and provide a quantitative basis for the conclusions regarding the improved accuracy and effectiveness of convolutional neural networks in shipping monitoring and vessel classification tasks.

Learning rate is one of the most important values. **Fig. 2.13** on the left shows the effects that occur when the SH changes.

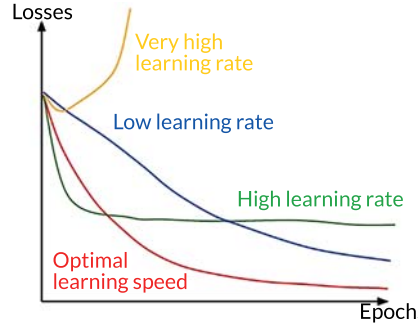


Fig. 2.13 Effect of learning rate on its accuracy

The second important thing to monitor is the accuracy of the network on the training and evaluation data. If to put them on the same graph, it is possible to assess the presence of overfitting, as evidenced by the diverging curves (**Fig. 2.14**).

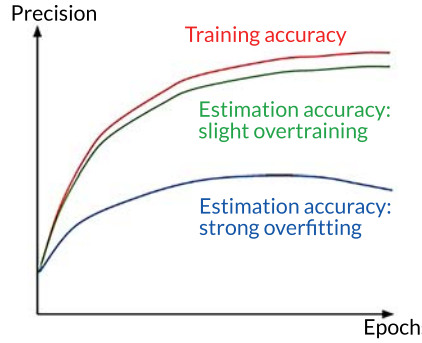


Fig. 2.14 Dependence of estimation accuracy on retraining

To find the optimal hyperparameters, it is possible to write a separate function that will independently select them and perform the optimization. In this case, it is better to use a non-uniform search (also known as "grid search"), but a random search, which often gives much better results.

2.4 Conclusion

The conducted research demonstrates that modern tasks of visual information recognition in water-area monitoring systems and inland waterway transport far exceed the capabilities of traditional computer vision algorithms. The analysis of theoretical and applied aspects of vessel-classification systems shows that methods based on heuristic rules, classical classifiers, or simple similarity metrics cannot ensure the required scalability or accuracy under real operating conditions of port infrastructure [21]. High variability in hull shapes, viewing angles, illumination, intraclass diversity, and the so-called *semantic gap* significantly complicate the construction of reliable models capable of generalizing information based on primitive features.

Based on theoretical justification and comparison of approaches, it has been established that only deep learning – and specifically convolutional neural networks (CNNs) – provides the necessary foundation for automatically extracting

relevant features from large sets of images [22]. CNN architectures demonstrate the ability to form multilayered representations, where initial layers respond to basic gradients and contours, and deeper layers extract complex spatial structures characteristic of particular vessel types. This approach ensures robustness to rotations, perspective changes, partial visibility, and various noise effects that are inevitable in video streams from river-port surveillance systems [23, 24].

Special attention was given to the training process of neural models, including the choice of loss functions, regularization techniques (l_1, l_2, elastic net), the use of batch normalization, optimization strategies, and the application of derivatives, gradient descent, and stochastic gradient descent. The findings show that training effectiveness largely depends on proper weight initialization, correct data preprocessing, and optimal hyperparameter tuning. The use of randomized search, cross-validation, and regularization significantly reduces overfitting and improves overall classification accuracy [25–27].

Thus, the results confirm that CNN-based models represent the most promising and effective technological foundation for developing intelligent monitoring systems for water transport. Their application provides:

– improved accuracy of vessel classification in port water areas;
– the capability for automated processing of large video streams in real time;
– adaptability to variable external conditions;
– reduced need for manual programming of rules and features;
– robustness to noise and incomplete data [28].

In the long term, implementation of the described models can ensure substantial progress in developing autonomous monitoring systems, enhance the efficiency of transport-flow management, optimize the operation of Ukrainian river ports, and strengthen safety mechanisms in the water-transport sector. At the same time, it can contribute to broader methodological application in areas such as organizational management, ecological analysis, engineering cybernetics, and simulation modeling, even including tasks related to the environmental protection of water resources [29]. The obtained results create a solid scientific basis for further advancement of intelligent image-analysis methods, integration of deep-learning models into complex surveillance systems, and development of new algorithms for vessel recognition in challenging environments.

Further research within the direction considered may focus on extending and refining the proposed approach. In particular, the use of continuous video streams instead of individual static images would allow temporal information to be incorporated and could improve classification reliability in complex scenarios involving object occlusions or visually ambiguous cases. Another important direction is the expansion

of the dataset to include images collected in different regions and under diverse operating conditions, which would enhance the generalization capability of the models.

A promising avenue for future work is the integration of visual data with additional information sources, such as navigation or sensor-based systems, in order to increase robustness and reduce the impact of noise and incomplete observations. Moreover, further studies may address optimization of convolutional neural network architectures with respect to computational complexity, inference latency, and hardware constraints, which are critical factors for practical deployment in real-time port monitoring systems.

Thus, the results obtained in this study provide a foundation for continued development of intelligent shipping monitoring systems and can serve as a starting point for expanding the functionality of decision-support and traffic management solutions for inland waterway transport.

Conflict of interest

The authors declare that there is no conflict of interest in relation to this paper, as well as the published research results, including the financial aspects of conducting the research, obtaining and using its results, as well as any non-financial personal relationships.

Financing

The study was performed without financial support.

Data availability

The data that support the findings of this study will be made available by the authors on reasonable request.

Use of artificial intelligence statement

The authors have used artificial intelligence technologies solely to verify the correctness and consistency of the English translation of the manuscript in an

academic style. In particular, the authors used the generative language model ChatGPT (OpenAI, GPT-5.1 Thinking) for language checking and minor stylistic refinement. The authors bear full responsibility for the content of the final manuscript; the AI tool is not credited as an author and is not responsible for the reported results.

Acknowledgments

The authors received no specific support from individuals or organizations that should be acknowledged beyond their institutional affiliations stated in the manuscript.

Authors' contributions

Pavlo Mykhalichenko: Conceptualization, Methodology, Supervision, Validation, Writing – review & editing.

Tetiana Cherniavska: Conceptualization, Theoretical framework, Formal analysis, Writing – original draft, Writing – review & editing.

Bohdan Cherniavskyi: Methodology, Data curation, Software, Investigation (experiments), Visualization, Writing – review & editing.

Viktor Nadtochii: Conceptualization, Domain expertise (water transport and monitoring systems), Validation, Formal analysis, Writing – review & editing.

Anatolii Nadtochyi: Methodology, Modeling and algorithmic development, Formal analysis, Interpretation of results, Writing – review & editing.

Maiia Korkh: Literature review, Investigation, Integration of results, Project administration, Writing – review & editing.

All authors contributed equally to the scientific content of this chapter, jointly developed the conceptual idea of the work, and approved the final version of the manuscript.

References

1. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S. et al. (2015). ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision, 115 (3), 211–252. https://doi.org/10.1007/s11263-015-0816-y

2. Krizhevsky, A., Sutskever, I., Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25, 1097–1105.
3. Er, M. J., Zhang, Y. (2023). Ship detection with deep learning: a survey. Artificial Intelligence Review, 56 (10), 11825-11865. https://doi.org/10.1007/s10462-023-10455-x
4. Zhang, B., Liu, J., Liu, R. W., Huang, Y. (2025). Deep-learning-empowered visual ship detection and tracking: Literature review and future direction. Engineering Applications of Artificial Intelligence, 141, 109754. https://doi.org/10.1016/j.engappai.2024.109754
5. Awais, C. M., Reggiannini, M., Moroni, D., Salerno, E. (2025). A Survey on SAR ship classification using Deep Learning. arXiv preprint arXiv:2503.11906. https://doi.org/10.48550/arXiv.2503.11906
6. Liu, D. (2023). TS2Anet: Ship detection network based on transformer. Journal of Sea Research, 195, 102415. https://doi.org/10.1016/j.seares.2023.102415
7. Wang, Y., Li, X. (2025). Ship-DETR: A Transformer-Based Model for EfficientShip Detection in Complex Maritime Environments. IEEE Access, 13, 66031–66039. https://doi.org/10.1109/access.2025.3559107
8. Galdelli, A., Narang, G., Pietrini, R., Zazzarini, M., Fiorani, A., Tassetti, A. N. (2025). Multimodal AI-enhanced ship detection for mapping fishing vessels and informing on suspicious activities. Pattern Recognition Letters, 191, 15–22. https://doi.org/10.1016/j.patrec.2025.02.022
9. Salem, M. H., Li, Y., Liu, Z., AbdelTawab, A. M. (2023). A Transfer Learning and Optimized CNN Based Maritime Vessel Classification System. Applied Sciences, 13 (3), 1912. https://doi.org/10.3390/app13031912
10. Agorku, G., Hernandez, S., Falquez, M., Poddar, S., Amankwah-Nkyi, K. (2024). Real-Time Barge Detection Using Traffic Cameras and Deep Learning on Inland Waterways. Transportation Research Record: Journal of the Transportation Research Board, 2679 (2), 703–720. https://doi.org/10.1177/03611981241263574
11. Simonyan, K., Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556. https://doi.org/10.48550/arXiv.1409.1556
12. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D. et al. (2015). Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1–9. https://doi.org/10.1109/cvpr.2015.7298594

13. He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 770–778. https://doi.org/10.1109/cvpr.2016.90
14. LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. Nature, 521 (7553), 436–444. https://doi.org/10.1038/nature14539
15. Lecun, Y., Bottou, L., Bengio, Y., Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86 (11), 2278–2324. https://doi.org/10.1109/5.726791
16. Goodfellow, I., Bengio, Y., Courville, A. (2016). Deep Learning. Cambridge: MIT Press, 775.
17. Ioffe, S., Szegedy, C. (2015). Batch normalization: accelerating deep network training by reducing internal covariate shift. Proceedings of the 32nd International Conference on Machine Learning, 448–456.
18. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. Journal of Machine Learning Research, 15, 1929–1958.
19. Kingma, D. P., Ba, J. (2014). Adam: a method for stochastic optimization. arXiv:1412.6980. https://doi.org/10.48550/arXiv.1412.6980
20. Hastie, T., Tibshirani, R., Friedman, J. (2009). The Elements of Statistical Learning. New York: Springer, 745.
21. Huang, I.-L., Lee, M.-C., Nieh, C.-Y., Huang, J.-C. (2023). Ship Classification Based on AIS Data and Machine Learning Methods. Electronics, 13 (1), 98. https://doi.org/10.3390/electronics13010098
22. Ren, S., He, K., Girshick, R., Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, 28, 91–99.
23. Leonidas, L. A., Jie, Y. (2021). Ship Classification Based on Improved Convolutional Neural Network Architecture for Intelligent Transport Systems. Information, 12 (8), 302. https://doi.org/10.3390/info12080302
24. Gallego, A.-J., Pertusa, A., Gil, P. (2018). Automatic Ship Classification from Optical Aerial Images with Convolutional Neural Networks. Remote Sensing, 10 (4), 511. https://doi.org/10.3390/rs10040511
25. Zhao, T., Wang, Y., Li, Z., Gao, Y., Chen, C., Feng, H., Zhao, Z. (2024). Ship Detection with Deep Learning in Optical Remote-Sensing Images: A Survey of Challenges and Advances. Remote Sensing, 16 (7), 1145. https://doi.org/10.3390/rs16071145
26. Redmon, J., Farhadi, A. (2017). YOLO9000: Better, Faster, Stronger. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, 6517–6525. https://doi.org/10.1109/cvpr.2017.690

27. Girshick, R. (2015). Fast R-CNN. 2015 IEEE International Conference on Computer Vision (ICCV). IEEE, 1440–1448. https://doi.org/10.1109/iccv.2015.169
28. Niranjan, A., Patial, S., Aryan, A., Mittal, A., Choudhury, T., Rabiei-Dastjerdi, H., Kumar, P. (2024). A Deep Learning Approach for Ship Detection Using Satellite Imagery. EAI Endorsed Transactions on Internet of Things, 10. https://doi.org/10.4108/eetiot.5435
29. Cherniavska, T., Nadtochii, V., Nadtochyi, A., Lomonosov, D., Cieślak, R., Cherniavskyi, B. et al.; Cherniavska, T. (Ed.) (2025). Organizational and structural modeling of the integration of marine robotics into multilevel environmental and ecological monitoring systems. Ecological systems modeling. Tallinn: Scientific Route OÜ, 169–195. https://doi.org/10.21303/978-9908-9706-6-0.ch8

CHAPTER 3

Reengineering of management processes for the restoration of transport and logistics infrastructure through image recognition and BIM-oriented remediation

Tetiana Cherniavska
Bohdan Cherniavskyi
Oksana Zghurska
Serhii Kasian
Kateryna Nakonechna
Yaroslava Mudra

Abstract

In the context of eliminating the consequences of emergency events (military conflicts, natural disasters, man-made accidents, etc.), it is appropriate to focus attention on issues of reengineering of infrastructure facilities, since the tasks go beyond traditional engineering design and include redesign of the managerial contour of the transport and logistics system (S&D → BIM/DT → 4D/5D → BPMN), re-assembly of roles and responsibility, implementation of end-to-end traceability (CDE/DT), risk-oriented prioritization and audit-ready frameworks of quality, safety and ecology. The authors of the monograph expanded the theoretical framework of the semantic content of the concept "reengineering" in the context of BIM/Digital Twin-oriented reconfiguration of the architecture and processes of remediation/reconstruction/restoration of transport and logistics infrastructure facilities on the basis of observation and diagnostics data. The increase in the scale and diversity of risks of a different nature determines the need to move from static regulations to a data-driven approach, namely: expanding the range of application of computer vision, analysis of images and UAV imagery, satellites, as well as neural-network recognition, which can be organically integrated with the BIM model of the facility, thereby forming a "digital twin". Such coupling will make it possible to provide a full cycle screening → diagnosis → prognosis → intervention, which will make it possible to automatically identify defects, verify their spatial-semantic localization in BIM, assess the degree of risk of collapse of the facility and/or failure, and on the basis of this – predict the operational life of structural elements. All this

 DOI: 10.21303/978-9908-845-05-0.ch3

will make it possible to carry out more accurate planning of remediation/reconstruction/restoration works (4D/5D), choose the optimal scenario and protocol of necessary measures, effectively manage the course of implementation of the complex of works for restoration with subsequent successful commissioning of the facility into operation and audit. The authors of the monograph substantiated that BIM technologies play the role of a driver in BPR (business process reengineering) of the transport safety management system as a whole, since they are able to combine surveillance and management of restoration and repair works within operational requirements in a single information contour, while increasing the speed, accuracy, safety and minimization of risks of different nature and scale. The study proposes a metric support, which will make it possible to assess the effect of the implemented complex of works. Thus, the result of reengineering is proposed to be assessed using an integral indicator BOR-Index, which includes an assessment of safety, time, cost, quality, DT-fidelity, completeness of evidence, timeliness, and the readiness of data and processes to be assessed using the I-Score index by levels of interoperability (syntax/formats, semantics, process, operational, evidence/CDE). The monograph studied and proposed a transferable benchmark of construction and restoration of the Genoa bridge (Italy) as confirmation of the feasibility of the coupling S&D → BIM/DT → 4D/5D in fast-track mode.

Keywords

BIM, Digital Twin, BIM-oriented remediation, restoration, transport safety, transport, logistics, infrastructure facility, interoperability, 4D/5D management, BPMN orchestration.

3.1 Introduction

First of all, attention should be focused on the relevance and significance of this problem area, which is determined by a number of facts, namely:

– the growth of the frequency and severity of emergency events, including military conflicts, which lead to large-scale damage to infrastructure, including transport and logistics infrastructure facilities, which directly affects the level of national security and carries the risk of economic losses;

– the limitations of the traditional approaches used for eliminating the consequences of emergency events and for the operational restoration of the transport and logistics system;

– the emergence of an innovative technological "window of opportunity", which is determined, first of all, by the scaling of the use of sensors, unmanned aerial

vehicles (drones), satellite data, edge-AI, as well as the maturity of BIM/IFC, and the introduction into the practice of carrying out repair and construction works of 4D/5D;

– the presence of a significant scientific-and-practical gap, which consists in the shift of the focus of scholars to the problems of defect detection of the functioning of transport and logistics infrastructure facilities and the not fully researched issues of using the potential of BIM/DT in the context of selecting optimal managerial decisions and quality control of remediation, reconstruction and restoration measures [1].

The authors of the study examined benchmarks of effective application of BIM, in particular, projects on eliminating the consequences of events of a military nature in Iraq [2] and the project of construction and restoration of an infrastructure facility in Italy [3], which updates the problem area of expanding the range of application of BIM/Digital Twin technologies in the transport and logistics sphere.

The authors defined the aim of the study – to form a holistic methodology and management architecture for the reengineering of affected transport and logistics infrastructure facilities based on end-to-end integration S&D → BIM/Digital Twin → 4D/5D+BPMN within the framework of ensuring transport safety, with a formally verifiable effect by the integral indicator BOR-Index, assessing safety, time, cost, quality, DT-fidelity, evidence, timeliness of the implemented complex of works and the interoperability profile I-Score [4, 5].

In the course of achieving the set goal, the authors performed the following tasks: substantiate the theoretical foundations of BIM-oriented remediation in the context of Surveillance & Diagnostics; form a conceptual model "surveillance → diagnostics → prognosis → intervention" in order to ensure operational management of transport safety; analyze the role and significance of the 4D/5D approach and at the same time to define the role of BIM as the "sense organ" of the digital twin; develop and propose an integration architecture and principles of interoperability; study the possibilities of transferring benchmarks of remediation/reconstruction/restoration of transport and logistics infrastructure facilities into projects for eliminating the consequences of emergency events.

3.2 The theoretical foundations of BIM-oriented remediation of transport infrastructure facilities

First of all, within the theoretical basis, it is necessary to define the key concepts and boundaries. The scientific reconnaissance of published sources on this topic carried out by the authors revealed that most works focus on detection [6]. However, in the authors' opinion, CV/ML can initiate the actions undertaken inside BIM/DT –

from a defect map to a work schedule and budgets. In such a case, it is necessary to substantiate semantically and describe the geo-referencing of the recognized object (in this case, an element of point and/or linear infrastructure). This concerns the transfer of a pixel mask/box into the object-oriented structure of BIM (for example, a bridge element, a road-surface layer, a section of railway track, etc.), which is critically important in the context of remediation, reconstruction and restoration as a result of repair-and-restoration works after emergency situations or military conflicts, but is rarely described as a logical bridge between AI and engineering practice [7].

The ideational impulse for searching for solutions to the tasks set in this study was a medical analogy. Namely, let's rely on the medical chain of continuity of care delivery → screening → diagnosis → prognosis → intervention → rehabilitation: screening as early identification of the condition, the clinical pathway through establishing a diagnosis and forecasting risk, and then therapeutic intervention and subsequent rehabilitation aimed at restoring functions and reducing disability. This intuitive sequence is used to clarify the logic of the engineering-and-managerial picture in the context of post-war and post-disaster restoration of transport and logistics infrastructure [8].

According to the authors' vision of the monographic study, architectural-and-process reengineering constitutes a purposeful, standards- and metrics-based transformation of data architecture, roles and services, as well as end-to-end managerial processes of the national (including regional) system for ensuring transport safety and its remediation in the process of eliminating the consequences of emergency events [9, 10]. This transformation covers the contour: "surveillance → image recognition → situation assessment → action planning → risk assessment → implementation of a complex of works on remediation, reconstruction and construction → commissioning of facilities into operation → post-monitoring" through integration of:

– Surveillance & Diagnostics (S&D), which includes computer vision/ML, UAVs, satellites (optical/SAR), (ground/airborne) LiDAR, IoT;

– BIM/Digital Twin (DT) as a semantic core (in accordance with IFC/BCF, ISO 19650), in which the states are fixed as-designed → as-built → as-damaged → as-repaired/as-remediated;

– the operational layer, which, among other things, includes BPMN orchestration, 4D/5D planning, as well as quality control, occupational safety, environmental control and audit.

Reengineering of the transport safety assurance system within the aggregate of all structural elements of the territorial transport and logistics system covers the entire multi-level ecosystem of participants (including public administrative bodies, operators, contractors, transport and communications regulators, financial donors, etc.), all classes of transport and logistics infrastructure facilities (linear and

point), the life cycle of remediation/reconstruction/restoration works, as well as the security and interoperability policy [11].

In the authors' conviction, the term "reengineering" most accurately describes the initial need to restore the functional properties of the transport and logistics system in the event of emergency situations, namely: destruction of a facility (complete or partial), disabling of infrastructure facilities and its unfitness for further operation, as well as complete and/or partial impossibility of carrying out the main processes for safety reasons. As a result of this, the relevant governing bodies face, determined by these circumstances, a goal – to radically redesign the architecture of data, roles, processes and decisions (S&D ↔ BIM/DT ↔ BPMN + 4D/5D) taking into account risks, standards and regulations, technical, technological, as well as financial capabilities (**Fig. 3.1**).

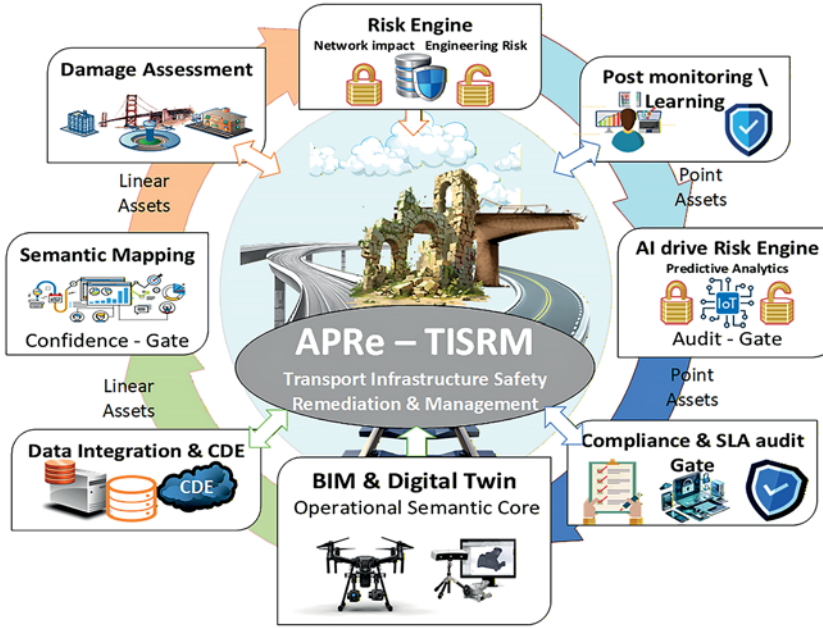


Fig. 3.1 Structure of APRe-TISRM (architectural-process reengineering for transport infrastructure safety, remediation & management/restoration)

Summarizing everything above, it is possible to state that BIM and Digital Twin form a single axis, where BIM acts as a semantic foundation (IFC/BCF), which makes it possible to structure the geometry, composition and hierarchy of all infrastructure elements,

the typology of defects and the necessary restoration-and-repair, as well as remediation/reconstruction/restoration works, also for the purposes of resource justification. Digital Twin acts as an operational "living" mirror above BIM, which makes it possible to synchronize the model with actual observations [12], as well as to record the states as-designed → as-built → as-damaged → as-repaired/as-remediated, implement scenarios and, while closing control loops, implement the function of a semantic container and a "sense organ" that is capable of integrating observation streams. In turn, Surveillance & Diagnostics is interpreted by the authors as a triad of interrelated functions: surveillance through the collection and processing of data of transport and logistics infrastructure facilities, recognition of patterns, and on this basis – diagnostics and prognosis with the corresponding engineering conclusions and managerial decision-making.

Surveillance & Diagnostics (S&D) act as a kind of "sense organ" of the system, which makes it possible to carry out comprehensive multimodal surveillance through the integration of the received information of video surveillance, UAVs, satellites, LiDAR, as well as IoT, etc. [13]. Then, identification and verification of patterns is implemented: the results of diagnostics and prognoses are semantically addressed to BIM elements or chainage segments (with an indication of the degree of confidence and the level of error).

According to the authors' view of the study, the uniqueness of BIM technologies lies in the fact that they directly influence critically important factors in the remediation/reconstruction/restoration of transport and logistics infrastructure facilities, namely, 4D (time) and 5D (cost) in operational management, scenario planning of the protocol of necessary works, resource optimization, risk reduction and ensuring safety [14]. 4D and 5D are layered onto the semantic model of organizing the implementation of a complex of measures that are multidirectional by nature, which makes it possible to implement optimal scenarios for restoring an infrastructure facility, while taking into account access possibilities, time windows, weather and technical-and-technological conditions, as well as the possibility of implementing combined measures. In turn, BPMN orchestration makes it possible to take into account and to highlight priorities in restoration scenarios into executable processes with human (Human-in-the-Loop) and automated multi-agent (AOAM) participation where required [15, 16].

Thus, in APRe-TISRM, BIM/DT acts simultaneously as a semantic container and an operational platform, which makes it possible to translate recognition results into manageable decisions, which as a result can significantly accelerate the remediation/reconstruction/restoration of facilities (4D), optimize the budget (5D), and at the same time ensure environmental and transport safety as a whole (including, through aggregating a database of quality control of performed works, compliance with standards and norms, etc.).

The next theoretically significant aspect is that, in the context of post-war and/or post-disaster recovery, the concept of remediation should be considered at least in terms of three interrelated layers [17], namely:

– *technical remediation (engineering remediation)*, which covers a set of measures to restore the main functional characteristics, eliminate defects and vulnerabilities, bring facilities to safety and reliability operation standards, and in some cases – redesign of transport and logistics hubs;

– *environmental remediation*, including, among other things, decontamination, extraction, sanitary-and-environmental control, returning the facility to economic use and preventing re-contamination;

– *socio-economic remediation*, namely, restoring accessibility and connectivity of the entire transport and logistics system, reducing the time of logistics costs, restoring network load, commodity flows, activating population mobility and employment, active use of services by all subjects of the system – that which in the logic of Build Back Better means not simply "return it as it was", but optimize the operational resource and increase the resilience of point and linear infrastructure facilities.

In accordance with what is indicated above, in APRe-TISRM this is reflected through the implementation of a complex of multidirectional remediation/reconstruction/restoration measures, covering a set of technical + environmental + socio-economic interventions, stitched through BIM/DT, making it possible to fix the semantics of the state of facilities, S&D (surveillance → recognition → diagnostics), 4D/5D scenario planning and BPMN orchestration with built-in Eco/Safety/Audit gates, as well as the corresponding KPIs.

Summarizing everything set out above, it is possible to state that BIM-oriented remediation is an integrated (technical, environmental and socio-economic) set of multidirectional interventions managed through BIM/Digital Twin, which play the role of a semantic-and-operational core, where S&D events are addressed to IFC elements and chainage segments, translated into 4D/5D scenarios with Eco/Safety gates, and the implementation is fixed as as-remediated in the CDE with an audit of the performed actions and calculated effects.

3.3 The methodological framework for managing remediation/reconstruction/restoration of affected transport and logistics infrastructure facilities

The process-and-architecture model APRe-TISRM substantiated in the previous chapter covers the end-to-end managerial contour of all remediation,

repair-and-restoration and construction measures for transport and logistics infrastructure facilities: from identification and verification of damage – to diagnostics of the possibility of putting into operation after carrying out the required protocol of works with account of prioritization, then to 4D/5D scenario planning in linkage to process orchestration of the set of measures, and at the final stage – to commissioning of the facility into operation on the basis of the conclusion of a comprehensive independent audit with subsequent post-restoration monitoring (with possible model training).

It should be especially emphasized that the semantic foundation of the framework sets the axis S&D → BIM → Digital Twin → BPMN + 4D/5D. In it, S&D (Surveillance & Diagnostics) plays the role of the "sense organ" of the system, where multimodal data sources, including information from video surveillance cameras, UAVs, satellites, LiDAR, etc., through fusion and registration are transformed into detections/masks/anomalies with an assessment of the degree of confidence. Aggregated S&D detections after fusion undergo geo-referencing and are addressed to the corresponding IFC elements for point infrastructure facilities and to chainage segments for linear infrastructure facilities with records of the degree of confidence and errors. All this is stored in the DT state and influences the Confidence-Gate/HIL.

BIM occupies a central place in this foundation; in it, data about a specific infrastructure facility are collected step by step and then, if necessary, used, starting from the design stage and ending with its demolition (including key technical characteristics, dimensions, cost, data on changes made during operation, etc.). Digital Twin serves as a kind of operational mirror, which makes it possible to synchronize the actual states of the facility (as-designed → as-built → as-damaged → as-repaired/as-remediated) and link the surveillance carried out during restoration with the chosen strategy, tactics, the established protocol, plans and deadlines of works. On this foundation, a risk engine of managerial decisions operates, which forms the prioritization of interventions and the choice of optimal scenarios of remediation/reconstruction/restoration of infrastructure facilities based on engineering risk, environmental and socio-economic risks, as well as the overall impact on transport safety.

The actual implementation of the functions of the framework described above is ensured by 4D/5D planning and BPMN orchestration. As indicated earlier, time (4D) and cost (5D) are layered onto the semantics of BIM in order to be able to form a set of permissible scenarios of remediation/reconstruction/restoration of facilities and choose from all permissible ones the optimal one, taking into account access windows, the level of safety of carrying out the protocol of measures, actually available resources and supply capabilities, as well as taking into account routing of construction-and-repair crews for linear infrastructure facilities and staging

of performed works for point facilities. The described processes can be presented as a chain of interrelated actions Intake → Verify (HIL) → Triage (risk) → → Plan (4D/5D) → Execute → QA/Accept → Monitor, where Human-in-the-Loop is engaged at a low degree of confidence of image recognition and in the case of critical extraordinary situations.

In the described configuration, the mechanism of manageability is regulated by "gates", namely: a Confidence-Gate for the purpose of verification of the segregation of S&D events with a low degree of confidence; an Eco-Gate in situations where carrying out a complex of remediation measures is mandatory from the standpoint of ensuring environmental safety (including remediation of the infrastructure facility itself, the territory on which it is located, as well as air and water) in cases of exceeding standards and regulations; a Safety-Gate in the case of the need to ensure access for carrying out the required protocol of works (including the need for demining, equalization the impact of weather conditions that hinder/make impossible the implementation of remediation/reconstruction/restoration operations, as well as taking into account actual technological windows); and at the final stage an Audit-Gate, which makes it possible to put the facility into operation only if there are artifacts as-repaired & as-remediated. All the above-mentioned artifacts, scenarios and protocols are maintained in the CDE based on the standards ISO 19650, IFC/BCF, BSDD, OGC/INSPIRE, which ensures interoperability, ensures automated interaction between heterogeneous systems without limitations, which makes it possible to carry out technical, donor financial and ESG audit.

It should be noted that, at the same time, the framework described above makes it possible to clearly distinguish two classes of infrastructure facilities – linear (including: highways, railways, tunnels, metro lines, etc.) and point (including: bridges, transport hubs, seaports, railway stations, airports, etc.). Each of them has different data addressing (chainage vs. IFC hierarchy), different profiled metrics, as well as different planning models and protocols for carrying out remediation/reconstruction/restoration works. It is important to note that this separation is embedded at all levels – from S&D mapping to KPI.

As was substantiated earlier in the study, in APRe-TISRM remediation is understood by the authors in an integrated manner, namely, as a set of various technical, environmental and socio-economic interventions in the logic of Build Back Better, in which the goal is not only to return the functional qualities of the facility, but also to optimize operational properties, increase its level of resilience, transport safety, and also increase the efficiency of the transport and logistics network as a whole. For these purposes, it is appropriate to implement in the framework a system of SLA/KPI management tools with the possibility of exporting them in open formats.

Thus, APRe-TISRM is not only a set of technologies, but also an executable methodology that makes it possible to transform the system of continuous monitoring of point and linear transport and logistics infrastructure facilities into a manageable system of their optimal operation based on the key parameters of ensuring transport safety, thanks to data standardization, process discipline and clearly defined responsibility.

Table 3.1 presented below serves as a bridge between the methodological description and the logic of the framework's functioning. It records those very "joints" of the data → decisions → actions → evidence contour, starting from the mapping of S&D events, including IFC elements and chainage segments – through updating the states of facilities in the Digital Twin – to launching the risk engine for managerial decision-making with the selection of strategy, tactics, priorities, optimal protocols and scenarios with subsequent 4D/5D planning, and then to process orchestration and completion of remediation/reconstruction/restoration works with an evidence base of compliance with norms and standards in the CDE.

Table 3.1 Method for calculating BIM/Digital Twin indicators in the APRe-TISRM framework

Indicators	Role in APRe-TISRM	Key variables	Formula/rule
1	2	3	4
Mapping of Pixel/Track objects → → IFC/Chainage	Transformation of S&D events into manageable data (IFC GUID/chainage segment)	$d \in D$ – an S&D event (detection/anomaly/change) obtained from surveillance sources; A; $u(d)$ – degree of confidence; $\varepsilon(d)$ – registration error	$f: D \to A$, where A is the set of infrastructure addresses: IFC elements (point assets) and chainage segments (linear assets); $f(d) = \text{argmax}_{a \in A}$ Overlap_{geo/sem} (d, a); $u(d)$ and $\varepsilon(d)$ are stored in the DT
Confidence-Gate	Filter for inaccurate data, as well as data with a low degree of confidence; $u(d)$ – confidence for event d [0...1]	$u(d)$ – confidence for event d [0...1]; $\varepsilon(d)$ – geo-registration error of event d (m/pixel/arc sec); θ_c, θ_ε – thresholds for the Confidence-Gate (minimum acceptable confidence and maximum acceptable error)	If $u(d) < \theta\_c$ or $\varepsilon(d) > \theta\_\varepsilon \Rightarrow$ HIL/inspection, prohibition of auto-generation of tasks until verification
Facility state → DT	Life cycle and history of a structural element of an infrastructure facility (versions in the CDE)	$x_i(t)$ – the state of DT element i over time (parameters of the detected defect, load/environmental indicators, statuses)	$x_i(t^+) = \text{Update}(x_i(t), f(d),$ evidence); states: as-designed → as-built → as-damaged → as-repaired/as-remediated

Continuation of Table 3.1

1	2	3	4
Engineering risk	Reliability/safety (probability × consequences)	P_i – probability of failure/degradation for element i (ML + physical models); S_i – consequences of failure (safety/lack of accessibility/downtime), damage scale	R_i^eng = $P_i \cdot S_i$; P_i – ML + physics; S_i – consequence matrices/risk classes
Eco-risk	Compliance with standards and regulations	c, A, T – contaminant concentration, impact area/volume, exposure time	R_i^eco = $\varphi(c, A, T)$ the specific φ is defined in the remediation methodology; "compliance vs exposure"
Network/node effect in transport and logistics infrastructure	Impact on the logistics corridor/transport hub	ΔL – network effect on a linear corridor (change in flow/travel time/detours/resilience); $\Delta Node$ – node effect (throughput of a transport hub, redundancy)	ΔL = g_net(Δflow, Δtravel-time, detours, resilience); $\Delta Node$ = g_node(throughput, redundancy/centrality)
Prioritization of remediation/reconstruction/restoration measures	Unified intervention priority	w_1, w_2, w_3 – weights for aggregating prioritization criteria (configured by the public administrative body/customer)	$\Pi_i = w_1 \cdot R_i$^eng + $w_2 \cdot R_i$^eco + + $w_3 \cdot$ (ΔL or $\Delta Node$); w is calibrated by the governing body; a multi-objective problem formulation is possible
Eco-Gate	Mandatory remediation when threshold contaminant pollution norms are exceeded	R_i^eco – environmental risk for element/segment i; θ_eco	If R_i^eco > θ_eco ⇒ remediation is mandatory (before and/or together with reconstruction/repair, or decommissioning); commissioning of the facility into operation is impossible without eco-evidence
4D/5D planning	Analysis of access to the remediation/reconstruction/restoration site; demining/technical-and-technological conditions/weather conditions/time windows	h(…) condition; θ_safe threshold	If access to the facility for executing the protocol of measures is not ensured ⇒ blocking the start of works
Execution of the work protocol → evidence base for control and audit	Closing tasks by evidence (CDE)	photos/videos, LiDAR/point clouds, environmental samples, as-built models, technical and financial documentation	Audit-Gate: closure ⇔ the CDE contains as-repaired & as-remediated with geo-referencing, timestamps, and links to the element/segment of the infrastructure facility

Continuation of Table 3.1

1	2	3	4
Post-monitoring/training	Self-improving loop	new events d'	Repeated surveys/sampling → updating x_i, retraining CV/ML, recalibrating Gate thresholds, updating maintenance plans for transport and logistics infrastructure facilities

The **Table 3.1** presented above is structured in such a way as to capture the key managerial invariants of APRe-TISRM:

a) what exactly enters at each stage (key variables);

b) which rule is applied, indicating the formula and/or specifying specific conditions;

c) the managerial effect that is achieved as a result of the actions undertaken.

Thus, BIM confirms its significance as a foundation, Digital Twin – as an operational "mirror" through synchronization of data on the state of facilities and the implementation of the selected protocol scenarios of remediation/reconstruction/restoration of transport and logistics infrastructure facilities, and the CDE – as an environment of interoperability, versioning and audit.

3.4 BIM/DT as the core of APRe-TISRM: verification of interoperability, significance for ensuring transport safety, and the BOR-Index

Within APRe-TISRM, which covers a set of interconnected and mutually influencing structural elements such as Surveillance & Diagnostics → BIM/Digital Twin → BPMN + 4D/5D → Evidence/CDE, it is precisely BIM/Digital Twin that act as the semantic and operational core. A critical condition for the operability of such a contour is interoperability, namely, the ability of heterogeneous systems and data, including IFC/BCF/BSDD, OGC/INSPIRE, CDE/API, etc., to pass the entire path from the source data to the decisions made and the evidentiary database confirming the implementation of the work protocol without losses and delays. It should be emphasized that losses during exchange and breaks at the junctions S&D ↔ BIM/DT ↔ BPMN/4D/5D significantly worsen safety, increase time and cost, and also negatively affect the completeness of evidence.

An in-depth study of this issue confirmed the primary role and significance of interoperability. In a general understanding, interoperability constitutes the ability

to automatically, safely and unhinderedly exchange interpretable data between devices and systems in an information-and-technology network in a standardized manner without technical limitations [18].

In the authors' opinion, assessing interoperability is one of the significant parameters of APRe-TISRM, since it makes it possible to connect the logic of the theoretical linkage S&D → BIM/DT → 4D/5D → CDE with practice, taking into account safety, time, cost, commissioning of infrastructure facilities into operation, which makes it possible to form an objective basis for ranking, financing and scaling BIM-oriented remediation for both linear and point facilities.

For the purpose of an objective assessment of interoperability, the authors emphasized the key aspects:

1. *Scope of coverage: linear vs point infrastructure facilities.*

As was indicated earlier, transport and logistics infrastructure covers linear facilities (including roads, rail tracks, subway lines, etc.) and point facilities (including bridges, tunnels, railway stations, transport hubs, logistics centers, terminals, seaports, airports, etc.) [19].

For linear facilities, the key aspects, in the monograph authors' view, are: chainage and dynamic segmentation, the continuity requirements of remediation/reconstruction/restoration works, routing of crews as well as equipment taking into account the availability of time windows, metrics of the load of the transport and logistics network, accessibility of facilities, and also the time and required resource provision for carrying out the necessary measures. In turn, for point infrastructure facilities, the depth of IFC hierarchies and domain MVDs, distribution by 4D/5D stages, and assessment of the criticality of load on a transport and logistics hub (throughput, redundancy, centrality) are of significant importance and role. In this regard, identification of risk profiles and the corresponding specific data necessitates separate verification and validation of interoperability for linear infrastructure facilities and separately for point infrastructure facilities.

In the context of this study, verification of interoperability is the answer to the question: "Do we correctly understand/interpret the obtained data about the infrastructure facility?". In turn, validation of interoperability is the answer to the question: "Do our integrated exchanges and processes S&D ↔ BIM/DT ↔ 4D/5D/BPMN actually work as intended and produce the required effect?" It should be emphasized that they are closely and end-to-end interrelated: without successful verification, valid effects are unstable and ultimately non-auditable, and without validation, the completed verification remains formal. However, together they transform BIM/DT from a "showcase" into an operational core for managing remediation and restoration for linear and point infrastructure facilities.

2. *Interoperability V&V framework (verification and validation).*

Four levels of interoperability and their corresponding checkpoints were examined in this study, namely:

– syntactic, which focuses on assessing the correctness of IFC/BCF/BSDD and OGC/INSPIRE formats and schemas, as well as exchange profiles/MVDs and validators;

– semantic, which consists in analyzing the consistency and comparability of defect types/codes/dictionaries for linear and point infrastructure facilities, as well as the accuracy of Pixel/Track → IFC/chainage mapping;

– process, oriented toward assessing consistency with BPMN orchestration, gate logic (including Confidence/Eco/Safety/Audit), statuses, as well as routes of remediation/reconstruction/restoration measures;

– operational, which focuses on assessing the causes and facts of delays, evaluating transaction reliability, and resilience under conditions of unstable communications and connectivity. In addition, the degree of completeness and traceability of evidence is assessed.

It is possible to note that for each level, the corresponding metrics are records: $\Delta Geom/\Delta Attr$ (round-trip losses), MapAcc (mapping accuracy), Latency (event → → DT → task), GatePass% (percentage of cases that passed the gates), Traceability (links evidence ↔ element/segment of the facility ↔ version), API-Reliability, CDE-Conformance. These indicators are aggregated into the I-Score (0…1) – an integral assessment of interoperability with calibratable weights and thresholds.

3. *Link between interoperability, safety, and the BOR-Index.*

In APRe-TISRM, it is possible to define safety as the primary priority (safety-first): through strict thresholds of the residual risk level at the level of elements and the required network accessibility/LOS, as well as through the Safety-Gate, namely through the possibility of safe access to the infrastructure facility, including by means of demining and carrying out remediation measures for soil, water and air, ensuring personnel-safe conditions for performing restoration works, as well as weather and technological conditions. In the monograph authors' view, interoperability is a predictor of safety (including transport safety), since it determines the accuracy and timeliness of defect mapping in the DT, the correctness of prioritizing the required remediation/reconstruction/restoration works, and the quality of the 4D/5D plan.

To objectify the assessment, it is proposed to use an integrated *BOR-Index (BIM-Oriented Remediation Index)* – an integral KPI of the quality of BIM-oriented remediation/restoration (0…1), which is based on normalized sub-indices, namely: the degree of reduction of engineering risk (S_1), eco-compliance (S_2), compliance with 4D parameters (S_3) and 5D parameters (S_4), DT consistency (S_5), completeness of evidence (S_6), and the degree of impact on the transport and logistics network (S_7).

Note that interoperability will affect the BOR-Index directly through S_5 (DT-fidelity) and S_6 (evidence completeness), and also indirectly through S_1/S_7 (assessments of risk accuracy as well as network accessibility). In practice, an increase in the I-Score leads to an increase in the BOR-Index, which is expressed in shorter delays in managerial decision-making, less rework, better manageability of 4D/5D, and faster auditable commissioning of restored infrastructure facilities into operation.

The mathematical formalization of the BOR-Index is as follows

$$BOR = \underbrace{G_{safe} \cdot G_{eco} \cdot G_{audit} \cdot G_{conf}}_{mandatory\ gates} \left(\sum_{k=1}^{7} \omega_k \cdot s_k \right), \sum_{k=1}^{7} \omega_k = 1, \omega_k \geq 0 . \tag{3.1}$$

The decomposition of the BOR-Index formula with gates (indicator functions) is presented below:

1. *Safety-Gate.*

$$G_{safe} = 1\left\{ \forall_i \in A : R_i^{res} \leq \bar{R}_i \right\} \cdot 1 \left\{ \begin{matrix} \forall_j \in W_{shed} : a_j \leq s_j \leq b_j - t_j \wedge Access_j = \\ = 1 \wedge Demining_j = 1 \wedge Wether_j = 1 \end{matrix} \right\}. \tag{3.2}$$

2. *Eco-Gate.*

$$G_{eco} = 1\left\{ \forall_i \in A_{eco} : R_i^{eco} \leq \Theta_{eco} \right\} \wedge 1\left\{ \forall_i \in A_{eco} : c_i \leq C_i^{\max} \wedge T_i \leq T_i^{\max} \wedge A_i \leq A_i^{\max} \right\}, \tag{3.3}$$

where $R_i^{eco} = \varphi(c_i, A_i, T_i); C_i^{\max}, T_i^{\max}, A_i^{\max}$ – approved limits based on standards/permits.

3. *Audit-Gate.*

Let's define the completeness and traceability of the evidence base in the CDE for all tasks being closed for remediation/reconstruction/restoration of transport and logistics infrastructure facilities

$$G_{audit} = 1\left\{ \forall_i \in W_{closed} : \underbrace{EvidComp_j \geq \tau_{evid}}_{\substack{completeness\ evidence \\ photo/video/LiDAR/eco\text{-}probes}} \wedge \underbrace{Trace_j = 1}_{\substack{connection\ to \\ IFG\text{-}GUID/chainage+ \\ +geo/time+signatures}} \wedge \underbrace{CDEconf_j \geq \tau_{cde}}_{\substack{ISO\ 19650:\ versions/statuses \\ WORM\text{-}logs}} \right\}. \tag{3.4}$$

4. *Confidence-Gate.*

$$G_{conf} = 1\left\{ \forall d \in D_{used} : \left(u(d) \geq \Theta_c \wedge \varepsilon(d) \leq \Theta_e \vee \mathrm{HIL}(d) = 1 \right) \right\}, \tag{3.5}$$

where $u(d)$ – the degree of confidence of the object detection; $\varepsilon(d)$ – the registration error; $\mathrm{HIL}(d) = 1$ – the fact of passing mandatory verification by specialists in the case of assessing "borderline" events.

The scheme presented in **Fig. 3.2** serves as a framework for linking the metrics described above. Thus, on the input contour S&D → CDE/DT, the I-Score is assessed for: Syntax/Formats and Semantics (registration and mapping accuracy); at the Prioritization/Planning junction – the I-Score for: Process/Operational (latency evt → DT → task, gate compliance); at the output Execute/QA → Post-Monitoring – the I-Score for: Evidence/CDE (completeness/traceability of evidence). All interoperability levels shown in the figure are aggregated into the BOR-Index, which makes it possible to directly identify the impact of each node of the scheme on the target effect of remediation and restoration of infrastructure facilities of the transport and logistics system.

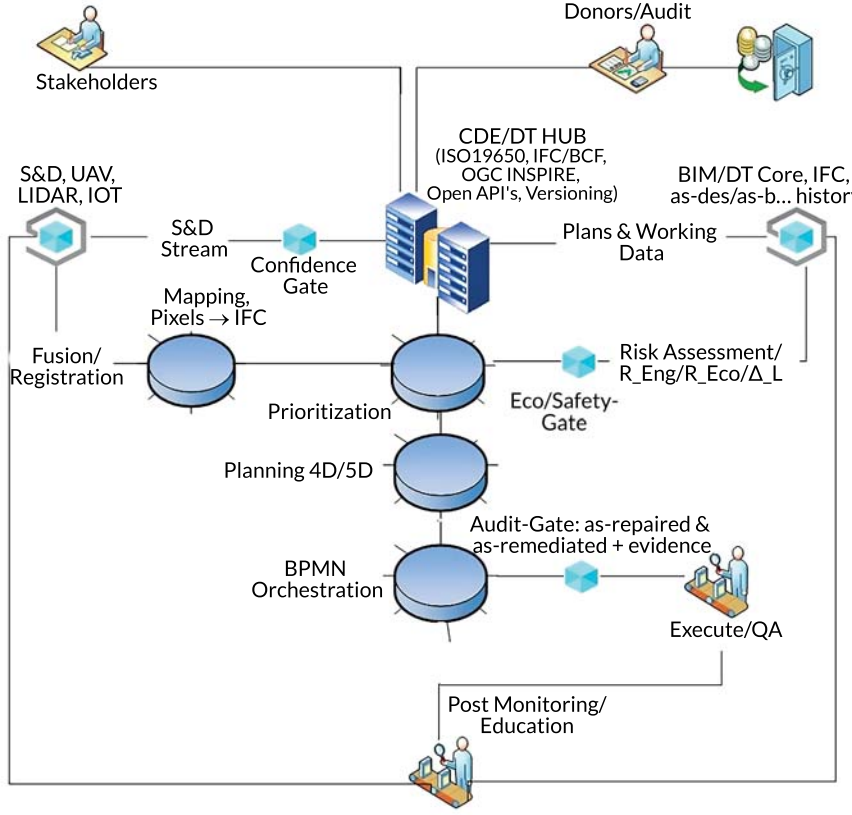


Fig. 3.2 Architectural map of APRe-TISRM flows

Next, let's move from the formal decomposition of the BOR-Index to its visual interpretation, since even a correct mathematical formalization remains a "black box" for various groups of stakeholders (let's refer to representatives of public administrative bodies, investors, IT specialists, engineers, contractors, etc.). Visualization makes it possible to interpret accumulated data and the calculated indicators in a clear way, namely:

1) it explicitly shows where exactly losses arise in the contours S&D → BIM/DT → → Ops (interoperability analysis by levels);

2) it explains which mechanism of improving interoperability is transformed into an increase in BOR (let's refer to the decomposition of cause-and-effect contributions);

3) it allows quick comparison of groups of infrastructure facilities (linear vs point) in terms of their weak/strong zones.

This is essential for the purposes of prioritizing remediation, reconstruction/restoration works, selecting the optimal scenario, protocol of measures, budget, as well as the timelines for its implementation.

Of all possible options, it is feasible to select the following visualization forms:

– *Heatmap*, which provides a "diagnosis" by interoperability levels (Syntax/Formats, Semantics, Process/BPMN + Gates, Operational, Evidence/CDE) and metrics (including Latency, MapAcc, Traceability), showing the specific breakpoints in the APRe-TISRM framework chain;

– *Waterfall*, which answers the question "Why does the BOR index grow?", i.e., it makes it possible to see the contribution of ΔS_5 (DT-fidelity), ΔS_6 (evidence completeness) and indirect $\Delta S_1/\Delta S_7$ (timeliness/accuracy), linking the driver (interoperability) and the result (BOR);

– *Radar (Linear vs Point)*, which makes it possible to compare profiles and see where linear transport and logistics infrastructure facilities fall behind point facilities (or vice versa).

In the monograph authors' view, together these three charts form a clear bridge "from formalization to making rational managerial decisions".

For calculations and charting, the authors used Python (Matplotlib) scripts in a Jupyter-type computing environment. This approach is reproducible (script versions and the source interoperability matrices are recorded), transparent for audit (all data transformations are documented), and vendor-neutral with respect to specific CDE/IFC validation vendors – normalized matrices/metrics agreed with the methodology presented above are provided as input. Numerical values in the charts are marked as reference design values and are subject to refinement in pilots.

In summary, it is possible to state that visualization in this context acts as a tool for managing the process as a whole, including managing information, coordinating all project participants, responsibility, resource provision, risks, time, cost, and the

results of remediation/reconstruction/restoration of transport and logistics infrastructure facilities. In the authors' conviction, this set of visualizations will make it easier to defend project decisions before public administrative bodies, investors, donors and contractors, while minimizing the level of semantic ambiguity.

The heatmaps presented in **Fig. 3.3** show where exactly the interoperability chain breaks, and not simply which class of transport and logistics infrastructure facilities is better or worse. Based on these data, a general conclusion can be made: for linear infrastructure facilities, the weak point is addressing, routing, and delays, and for point facilities – the consistency of the stages of remediation/reconstruction/restoration works and the KPI of transport hubs. These are the growth points of the BOR-Index.

By the aggregate interoperability indicator, point infrastructure facilities outperform linear ones, namely: I-Score_Point = 0.805 versus I-Score_Linear = 0.746. The weakest point for both groups is the operational delay "event → DT → task": the row average for Latency (inv.) for linear infrastructure facilities is ≈ 0.66 (values in the matrix respectively: 0.68/0.68/0.66/0.62/0.65), for point infrastructure facilities is ≈ 0.70 (values in the matrix respectively: 0.72/0.72/0.70/0.68/0.70).

The Waterfall (Linear vs Point) presented in **Fig. 3.4** shows which "building blocks" the increase in the BOR-Index is composed of due to the growth of interoperability. As is seen, the main contribution is made by ΔS_5 (DT-fidelity) and ΔS_6 (evidence completeness), and the indirect contribution – by $\Delta S_1/S_7$, which characterize timeliness and accuracy. For linear infrastructure facilities, the total increase amounts to ≈ +0.050 (= 0.030 + 0.015 + 0.005), for point infrastructure facilities – ≈ +0.060, i.e., with an equal improvement of interoperability, point transport and logistics infrastructure facilities obtain a slightly larger gain.

Based on this, it can be concluded that the maximum effect on the BOR index is produced by investment in the accuracy and relevance of the DT (reducing the gap "as-is" ↔ "DT") and the completeness of the evidence database in the CDE (traceable photos/videos/scans/environmental samples, etc.). For linear infrastructure facilities, it is additionally necessary to critically strengthen the chainage mapping chain and the VRPTW integration in order to reach the level of point infrastructure facilities. These waterfalls make it possible to answer the question: where to direct efforts first in order to transform I-Score growth into a tangible increase in the BOR-Index.

In the presented radar chart (**Fig. 3.5**), it is clearly visible that point infrastructure facilities consistently outperform linear transport and logistics infrastructure facilities across all five interoperability levels: the differences amount to ~0.04–0.06 on each axis (especially in Operational and Process (BPMN + Gates)). The value labels show: Linear ≈ 0.71–0.77, Point ≈ 0.77–0.82, i.e., the lag of linear infrastructure facilities is systemic rather than local.

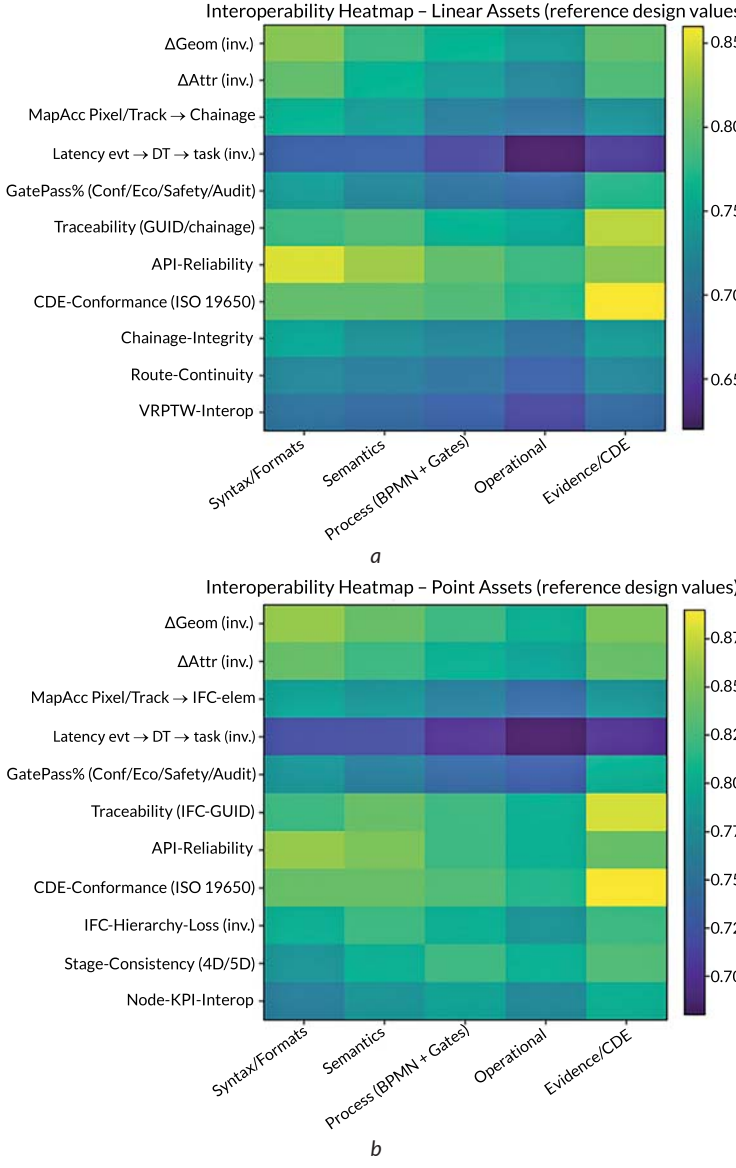


Fig. 3.3 Heatmaps of I-matrices for linear (*a*) and point (*b*) transport and logistics infrastructure facilities

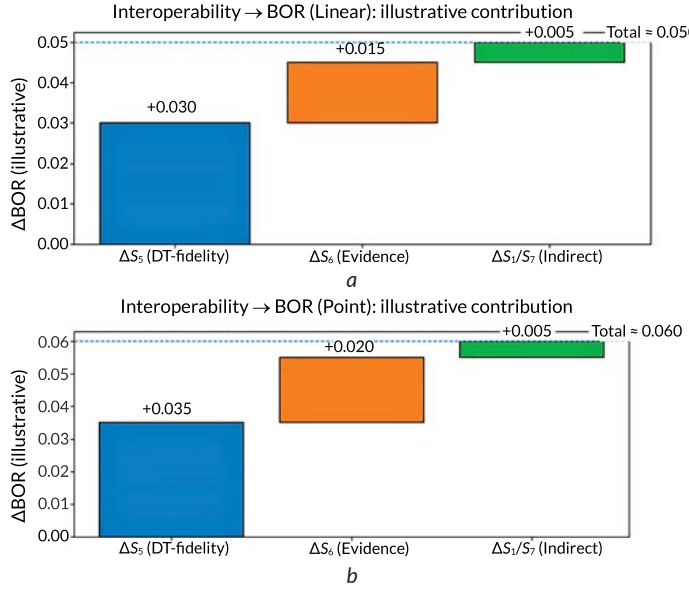


Fig. 3.4 "Waterfall" of effects: how interoperability increases BOR: *a* – linear transport and logistics infrastructure facilities; *b* – point facilities

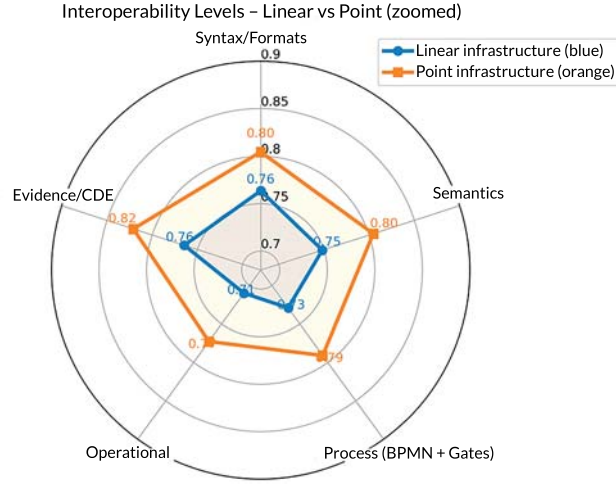


Fig. 3.5 Interoperability of S&D → BIM/DT → Ops by levels: comparison for point and linear transport and logistics infrastructure facilities

In practice, this means prioritizing the undertaken measures for linear infrastructure facilities precisely in the "middle of the contour", namely, reducing latency and improving semantic mapping/process gates, since this is where the most significant breaks for the BOR-Index are. It is important to note that the radar chart makes it possible to set a vector for a roadmap oriented toward rapid elimination of process-and-operational failures for linear infrastructure facilities, while maintaining the achieved high level of Evidence/CDE for point infrastructure facilities.

3.5 Benchmark case (Italy): transferable BIM/DT practices for remediation/reconstruction/restoration of infrastructure facilities in post-disaster territories

In order to substantiate the expediency of the BIM-oriented remediation approach proposed by the authors, the Genoa bridge restoration case (Italy) was analyzed in depth and comprehensively – a still rare, yet well-documented example where the linkage BIM ↔ Digital Twin ↔ 4D/5D ↔ evidence-based commissioning of an infrastructure facility actually worked under the conditions of a crisis restoration project. Implementation of this benchmark is absolutely expedient in the context of eliminating the consequences of natural and man-made disasters, as well as eliminating the consequences of military activity, since it:

– demonstrates a paradigm shift, namely, the transition from BIM as a design showcase to an operational core for managing time, cost, quality and safety;

– provides an evidence base for stakeholders, namely, real facts of applying the BIM/DT linkage in the urgent restoration of a critical infrastructure facility;

– directly correlates with the developed metrics, namely, this case makes it possible to test in practice the BOR-Index and I-Score proposed by the authors for implementation;

– covers validation of the group of point transport and logistics infrastructure facilities, since the Genoa bridge is a point asset with a high level of node criticality.

Description of the benchmark case: restoration of the bridge in Genoa (Viadotto Genova – San Giorgio, Italy).

Situation and scale. On 14 August 2018, a bridge span of approximately ≈ 210 m collapsed on the Morandi (Polcevera) viaduct, which as a result led to 43 fatalities, the shutdown of three railway lines, as well as significant transport-logistics and economic losses for Genoa itself and the region as a whole. A decision was made, as a replacement for the destroyed bridge, to build the new Genoa – Saint George

Bridge (Viadotto Genova – San Giorgio), which represents a mixed steel-and-concrete structure with a total length of 1,067 m and a width of approximately ≈ 30.8 m, supported by 18 reinforced-concrete piers. Sensor systems and robotic monitoring systems were incorporated into this infrastructure facility. The project cost amounted to about 202 million EUR. The opening of the facility took place on 3 August 2020, and the opening of traffic took place on 4 August 2020.

Technical and managerial parameters of the project. This project was implemented in an accelerated restoration mode with parallel dismantling of old structures, and the construction of the foundation started simultaneously with the completion of as-built design. Italferr (an engineering subsidiary company of Ferrovie dello Stato Italiane) was selected by the Pergenova consortium to develop the project. It should be noted that during design, 34 BIM models were synchronized into a single package. In order to meet extremely tight deadlines (initially, about three months were allocated for design), a BIM/DT approach with an open CDE environment on ProjectWise was implemented; MicroStation, OpenRoads, OpenBuildings Designer, Navigator, Descartes, and SYNCHRO (4D) were used. In addition, the geocontext of the data accumulated during project implementation was organically integrated, namely, LiDAR surveys of the terrain and orthophotos; digital surface model data of the relief and the base of the bridge structure for checking pile depths and aligning design decisions with actual elevations.

As a result, the following effects were obtained: optimization of the coordination process between project participants, automation of 4D scenarios, reduction of clashes, and increased accuracy of estimates and set planned indicators (**Fig. 3.6**).

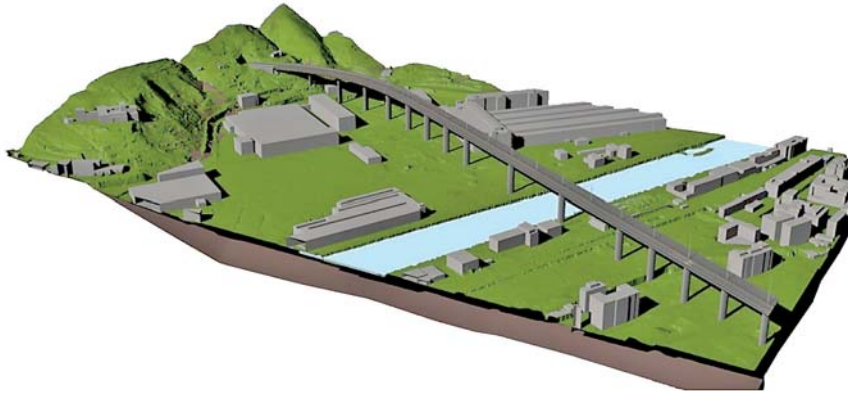


Fig. 3.6 The BIM information digital model for the Polcevera Viaduct
Source: [20]

Composition of participants and roles. Architectural concept – Renzo Piano; builders – Webuild and Fincantieri; operator/concessionaire – Autostrade per l'Italia; technical supervision and certification functions – RINA (a contract was signed for an amount of about 14 million EUR). Commissioning was carried out with the participation of regional government authorities. It should be noted separately that robotic inspection tools developed by the Istituto Italiano di Tecnologia were installed at the facility.

Chronology of the "compressed cycle". The public presentation of the concept took place in September 2018; intensive design-and-survey works → dismantling of the remaining structures in 2019; the "first stone" was laid on 25 June 2019; the main concreting of the deck slab was completed in June 2020.

As a result, intelligent 4D coordination of works on restoring the infrastructure facility made it possible to meet the target horizon of ≈ 15 months of construction.

Context and transferability boundaries in APRe-TISRM. In the Italian case, the linkage S&D → BIM/DT → Ops was built as a single managerial contour in which surveillance was immediately converted into manageable decisions. At the input, S&D provided multi-sensor surveying (UAV RGB/video, laser scanning, periodic inspections) with subsequent normalization and consolidation of data streams in a common CDE. In turn, at the BIM/DT layer, a high-detail model with specified LOD/LOI was maintained for critical nodes and in a mode of continuous synchronization of the states "as-designed ↔ as-built". The digital twin performed the role of a "sense organ" and a coordination plane for 4D/5D analysis with subsequent design and planning. Next, in the Ops layer, a 4D chronology of the executed works and 5D estimates were launched, and processes were orchestrated through BPMN with formalized Confidence/Eco/Safety/Audit gates. These gates served as "admission points", allowing the quality of data and the risks of starting/completing works within the specified time frames to be cross-aligned, with record of the evidence base in the CDE. As a result, an end-to-end cycle "data → semantics → action" was created and successfully tested, which made it possible to minimize time delays, significantly increase traceability, and ensure evidence-based commissioning of the facility into operation.

It should be especially emphasized that in the Italian project BIM/Digital Twin acted not simply as one of the data sources, but as the semantic and operational center of the entire contour S&D → BIM/DT → Ops. First, BIM/DT made it possible to ensure addressability and traceability. That is, each signal from S&D (data received from UAVs, LiDAR, change-detection, etc.) underwent normalization in the CDE and was then linked to IFC elements (via GUID, MVD validation), which helped eliminate hanging defects without geometric-and-semantic addressing. Second, the digital twin maintained continuous synchronization of the facility's states (as-designed → as-built → as-damaged → as-repaired/as-remediated) and served as a coordination

plane for 4D/5D, where each model update was automatically translated into the planned project timeline and budget. Third, BCF flows and BPMN gates in BIM/DT performed the role of a kind of process "arbiter". Finally, BIM/DT performed the role of the evidence base center: the "as-repaired & as-remediated" data were recorded in the CDE with linkage to the corresponding photo/video data, as well as data of technical expertise, environmental samples, and financial audit, which as a result made it possible to ensure the reproducibility of a direct link – "metric → artifact".

The case studied was interpreted in **Table 3.2** as a transferable correspondence table in terms of the parameters "data → tools → metrics → decisions", where each type of observation has a standard route in the CDE, semantic addressing in BIM/DT, and specific initial quality metrics.

Table 3.2 Generalized correspondence table relevant to APRe-TISRM

Block (data/tools)	Used sources and means	Processing → → integration	Target metrics/artifacts/ transfer characteristic
1	2	3	4
Geometry and condition	UAVs (RGB/video), ground/airborne LiDAR, photogrammetry, periodic inspections of the infrastructure facility	Registration of point clouds; alignment with BIM (as-built/as-damaged); upload to the CDE	ΔGeom, ΔAttr, scan density/coverage; deviation reports. Rule: each surveying cycle ↔ DT update; deviations automatically generate BCF issues
Semantics and addressing	IFC models, dictionaries, BCF issues	GUID traceability; mapping of S&D events → elements/ nodes of the infrastructure facility; MVD validation	MapAcc (Pixel/Track → → IFC), attribute completeness, % of valid GUIDs. Rule: no GUID – no task. Planning is performed only for events linked to IFC
Processes and integrations	CDE (in accordance with ISO 19650), BIM ↔ 4D/5D ↔ supply/ logistics API, starting BPMN engine	Auto-synchronization of statuses; orchestration of implemented measures and commissioning of the facility into operation; analysis of protocols of planned works	GatePass% (Conf/Eco/ Safety/Audit), status traceability, SLA compliance. Rule: gates as predicates: without Conf/Eco/Safety-OK status, tasks do not start; without Audit evidence, they do not close
Diagnostics and prognosis	Change detection, automated/semi-automated classification, engineering calculations of RUL/risk	Consolidation of conclusions in the DT; updating tolerance curves; task routing	Accuracy (F1/IoU), Latency evt → DT → task, RUL MAE/MAPE. Rule: from pixel to managerial decision: diagnostic conclusions automatically form 4D/5D tasks

Continuation of Table 3.2

1	2	3	4
Evidence and commissioning	Photos/videos, point clouds, certificates of completed works, environmental samples, etc.	Linking evidence to elements/nodes; "as-repaired/as-remediated" version in the CDE	Evidence completeness, audit trail, reproducibility. Rule: closure = evidence + version in the CDE: without this, the Audit-Gate does not pass
Operational readiness	4D/5D plans, supply/routing of construction-and-repair crews	Updating schedules as executed; VRPTW with account of access windows; accounting for Safety constraints	Timeliness (S_7), "plan vs actual" deviations, availability (uptime). Rule: latency down, uptime up: prioritize optimization of delays at the Ops stage

Thus, in the table presented above, each block of tools and data is closed on BIM/DT – from a pixel and a point cloud to a manageable IFC element, from the status of a structural element of an affected infrastructure facility as a result of emergency events to the plan of remediation/reconstruction/restoration works and to evidence, which is what makes possible the implementation of this case as transferable into APRe-TISRM.

3.6 Conclusion

As a result of the conducted study, APRe-TISRM was formed and substantiated – an architectural-and-process model of an end-to-end contour Surveillance & Diagnostics → BIM/Digital Twin → 4D/5D + BPMN → execution → post-monitoring, oriented toward restoration and remediation of transport and logistics infrastructure facilities. It was substantiated that BIM/DT is not a "showcase" of data of a linear or point facility, but performs the role of a semantic-and-operational core that links pixels/point clouds with manageable IFC elements and automatically translates updates of the facility state in linkage with 4D (time) and 5D (cost) of remediation/reconstruction/restoration. The study proposed the I-Score metrics for assessing interoperability and the BOR-Index as a target integral indicator of safety – time – cost – quality – evidence – timeliness of remediation/reconstruction/restoration works, providing a unified basis for data interpretation for all stakeholders. A comparative analysis of linear and point infrastructure facilities was carried out: for the former, the key failures are located in the chainage/operational latency chain (S_7), for the latter – in maintaining high DT accuracy and evidence completeness (S_5/S_6),

which determines the difference in priorities of implementing the work protocol. The Genoa bridge restoration case was studied in order to analyze the possibility of transferability of the practice of restoring an infrastructure facility located in post-disaster territories. This benchmark will be especially relevant and significant for countries that have suffered from emergency events of a natural and man-made nature, as well as from military activity. Thus, the aim of the study – to build a holistic methodology and management architecture for restoration and remediation based on image recognition and BIM/DT – was achieved, and the tasks of theoretical substantiation, formalization, metrics and benchmarking were completed. In the monograph authors' view, the APRe-TISRM model can be used as a pilot version for approbation on infrastructure facilities of Ukraine after the end of the active phase of hostilities, ensuring verifiable, accelerated and safe restoration of the national transport and logistics system.

The authors plan to conduct an in-depth study of the possibilities of 6D (operations) based on BIM/DT, where the digital twin will act as an operational core for managing energy consumption, operating modes and the technical condition of infrastructure facilities in real time. Research is also envisaged into algorithms of resilient operational management – MPC/RL – and the development of an extended BOR index, allowing quantitative assessment of the effect of 6D interoperability through reduction of OPEX/CO_2e, increase of uptime, and improvement of safety.

Conflict of interest

The authors declares that there is no conflict of interest regarding the content of this paper, including any financial aspects related to the conduct of the research, the acquisition and use of its results, or any non-financial personal relationships.

Financing

The study was performed without financial support.

Data availability

The data that supports the findings of this study will be made available by the authors on reasonable request.

Use of artificial intelligence statement

The authors state that artificial intelligence was used in this article only for partial translation of the text and correction of academic style.

Acknowledgments

The authors received no specific support from individuals or organizations that should be acknowledged beyond their institutional affiliations stated in the manuscript.

Authors' contributions

Tetiana Cherniavska: Conceptualization, Development of the theoretical and methodological framework (APRe-TISRM), Formal analysis, Integration of BIM/Digital Twin and Surveillance & Diagnostics concepts, Writing – original draft, Writing – review & editing.

Bohdan Cherniavskyi: Methodology, Data curation, Software and computational modeling, Image recognition and analytics (CV/ML), Visualization (heatmaps, indices, benchmarking diagrams), Writing – review & editing.

Oksana Zghurska: Investigation, Literature review, Analysis of post-conflict and post-disaster infrastructure recovery practices, Interpretation of results, Writing – review & editing.

Serhii Kasian: Domain expertise in transport and logistics infrastructure, Validation of engineering assumptions, Risk and safety analysis, Formal analysis, Writing – review & editing.

Kateryna Nakonechna: Methodological support, Analysis of process reengineering (BPMN, 4D/5D planning), Integration of management and operational aspects, Writing – review & editing.

Yaroslava Mudra: Case study analysis (benchmarking, including international experience), Data interpretation, Project administration, Coordination of interdisciplinary inputs, Writing – review & editing.

All authors contributed to the scientific content of this study, jointly developed the conceptual idea of BIM-oriented remediation and process reengineering for transport infrastructure recovery, participated in discussions and validation of results, and approved the final version of the manuscript.

References

1. Abdelalim, A. M., Essawy, A., Sherif, A., Salem, M., Al-Adwani, M., Abdullah, M. S. (2025). Optimizing Facilities Management Through Artificial Intelligence and Digital Twin Technology in Mega-Facilities. Sustainability, 17 (5), 1826. https://doi.org/10.3390/su17051826
2. Saeed, Z. O., Almukhtar, A., Abanda, H., Tah, J. (2021). BIM Applications in Post-Conflict Contexts: The Reconstruction of Mosul City. Buildings, 11 (8), 351. https://doi.org/10.3390/buildings11080351
3. Adopting Digital Twin Solutions for Genoa Bridge Reconstruction Proves to be Economical and Efficient. Geospatial World Prime. Available at: https://geospatialworld.net/prime/case-study/aec/adopting-digital-twin-solutions-for-genoa-bridge-reconstruction-proves-to-be-economical-and-efficient/
4. Gordo-Gregorio, P., Alavi, H., Forcada, N. (2025). Decoding BIM Challenges in Facility Management Areas: A Stakeholders' Perspective. Buildings, 15 (5), 811. https://doi.org/10.3390/buildings15050811
5. Cherniavskyi, B.; Cherniavska, T. (Ed.) (2025). Digitalization of crisis management remediation: assessment of implementation and development prospects. Economy in the era of digital transformation: trends, opportunities and perspectives. Tallinn: Scientific Route OÜ, 51–73. https://doi.org/10.21303/978-9908-9706-0-8.ch3
6. Wu, D., Zheng, A., Yu, W., Cao, H., Ling, Q., Liu, J., Zhou, D. (2025). Digital Twin Technology in Transportation Infrastructure: A Comprehensive Survey of Current Applications, Challenges, and Future Directions. Applied Sciences, 15 (4), 1911. https://doi.org/10.3390/app15041911
7. Cherniavskyi, B.; Cherniavska, T. (Ed.) (2025). Information technologies in scenario-based modeling of post-conflict territory remediation: from express sanitation to sustainable recovery. Ecological systems modeling. Tallinn: Scientific Route OÜ, 74–95. https://doi.org/10.21303/978-9908-9706-6-0.ch4
8. Cherniavskyi, B., Blakyta, H., Susidenko, V., Andreichenko, A., Remyha, Y., Podmazko, O.; Cherniavska, T. (Ed.) (2025). Innovative technologies and digital models in the post-war recovery of the transport and logistics system of Ukraine. Economy in the Era of Digital Transformation: Trends, Opportunities and Perspectives. Tallinn: Scientific Route OÜ, 110–143. https://doi.org/10.21303/978-9908-9706-0-8.ch5
9. Elazouni, A., Salem, O. A. (2011). Progress monitoring of construction projects using pattern recognition techniques. Construction Management and Economics, 29 (4), 355–370. https://doi.org/10.1080/01446193.2011.554846

10. Remedial Action Work Plan (2022). Hunter Army Airfield, G.
11. Shen, C., Chou, C.-C. (2010). Business process re-engineering in the logistics industry: a study of implementation, success factors, and performance. Enterprise Information Systems, 4 (1), 61–78. https://doi.org/10.1080/17517570903154567
12. Cherniavska, T., Cherniavskyi, B. (2025). Digital reconstructor: Integration of digital twins for the reconstruction and remediation of war-affected territories in Ukraine. UKLO Proceedings, 1 (1), 137–145. https://doi.org/10.20544/AISC.1.1.25.P13
13. Seidaliyeva, U., Ilipbayeva, L., Utebayeva, D., Smailov, N., Matson, E. T., Tashtay, Y. et al. (2025). LiDAR Technology for UAV Detection: From Fundamentals and Operational Principles to Advanced Detection and Classification Techniques. Sensors, 25 (9), 2757. https://doi.org/10.3390/s25092757
14. Hosamo, H. H., Rolfsen, C. N., Zeka, F., Sandbeck, S., Said, S., Sætre, M. A. (2024). Navigating the Adoption of 5D Building Information Modeling: Insights from Norway. Infrastructures, 9 (4), 75. https://doi.org/10.3390/infrastructures9040075
15. Tebourbi, H., Nouzri, S., Mualla, Y., El Fatimi, M., Najjar, A., Abbas-Turki, A., Dridi, M. (2025). BPMN-Based Design of Multi-Agent Systems: Personalized Language Learning Workflow Automation with RAG-Enhanced Knowledge Access. Information, 16 (9), 809. https://doi.org/10.3390/info16090809
16. Cherniavska, T., Cherniavskyi, B. (2024). Architecture-oriented agent-based model (AOAM) for optimizing transport evacuation management and emergency medical assistance in the context of the war in Ukraine: challenges and prospects. CEUR Workshop Proceedings. Available at: https://ceur-ws.org/Vol-3892/paper21.pdf
17. Harclerode, M., Ridsdale, D. R., Darmendrail, D., Bardos, P., Alexandrescu, F., Nathanail, P. et al. (2015). Integrating the Social Dimension in Remediation Decision-Making: State of the Practice and Way Forward. Remediation Journal, 26 (1), 11–42. https://doi.org/10.1002/rem.21447
18. Berg, C. (2024). Interoperability. Internet Policy Review, 13 (2). https://doi.org/10.14763/2024.2.1749
19. Gomes Correia, A., Winter, M. G., Puppala, A. J. (2016). A review of sustainable approaches in transport infrastructure geotechnics. Transportation Geotechnics, 7, 21–28. https://doi.org/10.1016/j.trgeo.2016.03.003
20. The new Genova San Giorgio Viaduct. Elegance, Innovation, Resilience. Implemented Solutions (2022). Gruppo FS Italiane. Available at: https://www.italferr.it/en/technology-and-innovation/BIM/design/genoa-san-giorgio-viaduct.html

CHAPTER 4

From pattern recognition to remediation management in a closed digital loop architecture for post-war logistics

Tetiana Cherniavska
Bohdan Cherniavskyi
Alla Rusnak
Iryna Nadtochii
Artur Harahulia
Oleksii Bobrovskyi

Abstract

The chapter provides an in-depth study of the problem of transition from digital diagnostics to executable management in the tasks of implementing remediation processes in areas affected by emergency events, including those of a military nature, where the effectiveness of recovery measures is largely determined by the quality of logistical coordination, the accessibility of infrastructure facilities, and the speed of managerial decision-making. Remediation in this study is interpreted as a complex, multidimensional process that includes the elimination of the consequences of war-related destruction and various kinds of contamination of soil, water, and air, the restoration of the spatial connectedness of territories, the provision of safe access, as well as the implementation of environmental, infrastructural, and socio-economic measures of post-crisis recovery. Against this background, the necessity is substantiated of transitioning from the fragmented use of digital tools to an integrated management architecture ensuring a closed cycle of "observation – recognition – geospatial diagnostics – decision – execution – verification – adaptation". As a conceptual foundation, the architecture of the Closed Digital Diagnostic Loop for Remediation Logistics is proposed, in which the Pattern Recognition block is not an autonomous analytical module, but acts as a source of diagnostic events forming inputs for geospatial integration, decision support systems, and simultaneous process orchestration. The central element of the architecture is the Decision & Orchestration Core (DSS (Decision Support System) + BPMN (Business Process Model and Notation)/CPM (Critical Path Method)) linkage, which ensures the transformation

 DOI: 10.21303/978-9908-845-05-0.ch4

of the results of recognition and geospatial diagnostics into executable managerial actions in the logistics of remediation operations, including, among other things, task prioritization, resource allocation, as well as schedule planning and control of the sequence of the work being performed. Separately, the monograph reveals the role of AR/VR (Augmented Reality/Virtual Reality) and Human-in-the-Loop mechanisms as an end-to-end HMI layer, which ensure the interpretation of diagnostic results, the coordination of decisions between levels of management, as well as support for the field implementation and verification of measures.

This study has a conceptual-methodological character and is aimed at substantiating the architectural integration of observation technologies, pattern recognition, GIS (Geographic Information System), DSS, BPMN/CPM, and HMI (Human – Machine Interface) mechanisms within a unified diagnostic-managerial loop for the implementation of a complex of multidirectional tasks of post-war remediation logistics. The proposed authorial concept creates a methodological basis for the further formalization of the rules of transition from diagnostic events to managerial actions, the development of domain orchestration profiles, and subsequent scenario-based as well as empirical validation under conditions of post-crisis recovery. The monograph considers transport and logistics as the central applied profile, because it is precisely there that the diagnostics → orchestration → execution linkage is most critical, and also because they are critical subsystems through which the managerial mechanism for eliminating the consequences of emergency events and subsequent remediation is implemented. However, the concept substantiated by the authors has a broader scope of applicability, namely: the elimination of emergency events not only of a post-conflict, but also of a natural as well as technogenic nature in all spheres of the socio-economic system, and the remediation of territories as a whole.

Keywords

Pattern recognition, remediation logistics, post-war remediation, emergency situations (ES), closed digital diagnostic-managerial loop, observation systems, geographic information systems, decision support systems, BPMN, critical path method, AR/VR, Human-in-the-Loop.

4.1 Introduction

It should be noted that the unprecedented scale of destruction and the duration of military activity of a hybrid nature in Ukraine have caused systemic destruction, affected and destructively influenced not only the physical infrastructure, but also the ecological condition of territories, production chains, the socio-economic

connectedness of regions and, as a consequence, the possibilities of rapid recovery. As the damage increases, it becomes obvious that post-war recovery cannot be considered exclusively as a construction and/or engineering task: it requires an integrated managerial approach capable of linking into a single whole the diagnostics of damage of various kinds, risk assessment and forecasting, the algorithmization of remediation and recovery works, their spatial, technical, and resource planning, as well as the logistical coordination and control of the execution of adopted managerial decisions.

In this context, the remediation of affected territories acquires special significance, understood by the authors as a set of measures differing in the nature of execution, aimed at eliminating the consequences of war-related destruction and contamination, restoring safe access to objects and territories, ecological stabilization, and creating conditions for socio-economic reintegration.

In the monographic study, the focus of attention is concentrated on transport and logistics; however, the conceptual idea of implementing remediation on the basis of an integrated management architecture ensuring a closed cycle of "observation – recognition – geospatial diagnostics – decision – execution – verification – adaptation" is absolutely applicable to other sectors of the national economy as well. In scientific publications devoted to the topics of humanitarian logistics and resilience, it is consistently emphasized that the effectiveness of restoration and remediation is largely determined by the quality of logistics coordination, digital integration, and the adaptability of decisions [1–4].

Thus, according to the authors' conviction, transport and logistics in this system simultaneously perform two roles at once: on the one hand, as an object of recovery (including, among other things, the necessity of reconstruction and remediation of destroyed roads, bridges, hubs, seaports and river ports, airports, railway network facilities) and, on the other hand, as an instrument for the implementation of remediation measures (in this case, this refers to the delivery of resources, the mobilization of equipment, ensuring access to intervention zones, supporting interagency coordination at all levels, etc.). It is precisely for this reason that the transport-logistics system is justifiably considered in this monograph as a critical contour of post-war recovery, on the effectiveness of which depend the speed of work execution, the safety of operations, and the attainability of the target results of remediation, reconstruction, and recovery. Along with this, one should take into account the fact that under conditions of resource scarcity, as well as the spatial heterogeneity of damage, taking into consideration their different character and the high dynamics of the situation, the key factor becomes not only the availability of data on the condition of objects and territories, but also the ability to quickly transform these data into executable managerial actions.

The relevance of such a conceptual approach is also confirmed by the scale of the losses accumulated since the beginning of the war. Thus, according to the updated joint assessments of the Ukrainian government, the World Bank, the European Commission, and the UN (RDNA), the total cost of the recovery and reconstruction of Ukraine continues to grow: the estimate amounting to 486 billion USD in RDNA3 (as of the end of 2023) was increased to 524 billion USD in RDNA4 (as of the end of 2024), and in the update published on February 23, 2026 (RDNA5, covering the period up to December 2025), the aggregate needs for remediation, reconstruction, and recovery are already estimated at approximately 588 billion USD, while direct damage exceeds 195 billion USD. Among the most affected sectors are the housing sector, transport, and energy [5–9]. All of the above-listed data indicate that post-war recovery cannot be built on point-based, fragmented solutions; the foundation of such recovery must be an architecturally organized digital management loop that will make it possible to ensure prioritization and adaptive reallocation of resources under conditions of an updated situation.

Despite the active development of digital tools for monitoring, damage mapping, and decision support, in the practice of eliminating the consequences of emergencies and post-crisis remediation, a substantial gap remains between the stages of observation, diagnostics, and the execution of managerial decisions. Observation systems (including UAVs (Unmanned Aerial Vehicle), satellite data, thermal and multispectral imaging) are capable of ensuring the operational collection of information; however, by themselves they do not form managerial actions. Methods of pattern recognition and intelligent analytics make it possible to identify damage, risk zones, and changes of condition, but their results often remain at the level of cartographic or reporting interpretation [10, 11]. GIS and DSS strengthen the spatial and analytical component; however, without formalized process orchestration and mechanisms for execution control, they do not ensure the completeness of the managerial cycle. As a result, the managerial loop proves to be broken, while the speed and quality of decisions are reduced precisely under those conditions where the cost of error and delay is maximally high.

In the present study, a conceptual-methodological solution to the indicated problem is proposed on the basis of the architecture of a closed digital diagnostic-managerial loop for the remediation logistics of post-war recovery. The key idea consists in the transition from pattern recognition to remediation management through a formalized mechanism for translating diagnostic events into executable tasks, scenarios, and processes. In contrast to approaches in which the Pattern Recognition block is considered as an autonomous analytical module, in this work it is conceptualized as a source of diagnostic events and confidence/uncertainty assessments, integrated into the geospatial layer and further transformed into managerial actions.

The main element of the proposed architecture is the Decision & Orchestration Core (DSS + BPMN/CPM), ensuring the prioritization of remediation tasks, the planning of resources and timeframes, the formalization of the sequence of operations, and execution control under conditions of a changing situation. Such a core makes it possible to connect digital diagnostics with the process implementation of decisions, thereby forming an executable managerial loop rather than a set of disparate digital tools. Additionally, the study expands the role of AR/VR and Human-in-the-Loop mechanisms, which are considered not as auxiliary visualization, but as an end-to-end HMI layer of interpretation, coordination, field support, and verification of decisions at all phases of the loop.

It should be emphasized that the proposed architecture has a broader horizon of applicability and may be used for the tasks of eliminating the consequences of emergency situations of a natural, technogenic, and military nature, as well as for the post-crisis remediation of territories in various domains. At the same time, in this work the transport-logistics sphere is considered as the priority implementation profile, since it is precisely in it that the necessity of a rapid and reliable transition from diagnostic data to managerial action is manifested most clearly [12]. This makes it possible simultaneously to preserve the interdisciplinary scale of the concept and to ensure its subject-specific focus on tasks of high practical significance.

The purpose of the study is to substantiate and structure the concept of the Closed Digital Diagnostic Loop for Remediation Logistics, in which the results of observation and pattern recognition are transformed into executable managerial decisions through the integration of GIS, DSS, BPMN/CPM, and an end-to-end HMI layer. Within the framework of this purpose, emphasis is placed on the architectural integrity of the loop, the logic of the transition from diagnostics to process orchestration, as well as on determining the directions and prospects of practical application [13, 14].

To achieve the stated purpose, the following tasks are solved in the work: clarification of the subject field of research at the intersection of the elimination of the consequences of emergency situations (of a natural, technogenic, and military nature), post-crisis recovery, and remediation logistics, with substantiation of the transport-logistics sphere as the priority domain for implementation of the concept; substantiation of the necessity for and development of the conceptual architecture of a closed digital diagnostic-managerial loop (Closed Digital Diagnostic Loop) for post-war remediation logistics, including interrelated functional blocks of observation, recognition, geospatial integration, decision orchestration, execution monitoring, and feedback; determination of the role of the Pattern Recognition block and geospatial diagnostics as a source of diagnostic events (including confidence/uncertainty assessments) forming inputs for the managerial core and

ensuring the transition from data analysis to decision-making; disclosure of the role of AR/VR and Human-in-the-Loop mechanisms as an end-to-end HMI layer and determination of the prospects for scaling the proposed architecture to a wider range of tasks of eliminating the consequences of emergency situations and the remediation of territories, including the formation of domain profiles for further research [14].

4.2 Degree of study of the problem and research gaps in the field of digital diagnostics and management of remediation logistics

First of all, it should be noted that the problematics of digital management of the elimination of the consequences of emergency situations (ES) and post-crisis recovery have been developing rapidly in recent years; however, predominantly along separate directions: process management, geographic information systems, artificial intelligence, decision support systems, platform-based data integration, and visualization. The scientific exploration by the authors of the problematics under study revealed the following. Thus, such aspects as the following are covered in open sources quite deeply and broadly:

– the use of remote sensing, satellite data/UAV for crisis monitoring and damage mapping. For the tasks of recognition/segmentation/classification in the post-disaster environment, there already exist high-quality datasets and an established scientific base. In particular, RescueNet offers high-resolution UAV images with annotations for classification and semantic segmentation, specifically oriented toward damage assessment after disasters [15]. This constitutes a weighty argument in favor of the fact that the transition to pattern recognition proposed in this study is not an artificial "expansion of the topic", but relies on the already established empirical and methodological base of CV/remote sensing for crisis tasks;

– the application of deep learning/pattern recognition for the purposes of detection, segmentation, damage assessment, and change detection in the disaster context;

– the use of AI in disaster management. Contemporary reviews show that AI is applied for forecasting, risk assessment, evacuation planning, resource allocation, and big data analysis, but also emphasize the problems of trustworthiness, bias, explainability, and integration into real decision-making loops [16]. This creates a strong foundation for the Pattern Recognition & Diagnostic Analytics block in your architecture, but simultaneously indicates the necessity of a separate mechanism for transforming recognized patterns into managerial actions;

– the application of GIS/Web-GIS platforms as a basis for integrating observations and modeling for the purpose of making effective decisions. Thus, a separate

cluster of studies on GIS and WebGIS in risk and emergency management is well developed. For example, the RiverCure Portal demonstrates that a Web-GIS platform can combine observations, modeling, the work of different organizations, and decision support at various stages of the risk management cycle (preparedness, operational response, risk assessment), including data assimilation and a multi-stakeholder configuration [17]. This is especially valuable for our concept, because geospatial integration in it must perform not a decorative, but a diagnostic-operational function (risk layers, accessibility, constraints);

– the advantages of BPMN and workflow approaches for formalizing emergency/disaster response processes [18]. Thus, in the works of Betke and co-authors, methodological extensions of BPMN for disaster response processes are discussed, which confirms the scientific viability of process modeling in the crisis sphere;

– the features of adaptive workflow execution in emergency response (including for run-time reconfiguration, execution engines);

– CPM in individual tasks of crisis management, including the elimination of the consequences of emergency situations, as well as the planning of remediation operations and emergency logistics. Thus, CPM has long been established in project management as a tool for identifying the critical sequence of works, task dependencies, and schedule control. In modern interpretations, its role as a communication and management tool for complex projects, and not only a computational mechanism, is emphasized [19]. For remediation logistics, this is directly relevant, since time delays at critical nodes (routes, access to facilities, infrastructure recovery) rapidly scale risks and cost;

– the application of integrated platforms (BIM/GIS/IoT/monitoring) and visualization in humanitarian logistics and crisis management for situational awareness and support of managerial decisions. Contemporary platform solutions already demonstrate the value of integrating GIS, BIM, and real-time monitoring for multi-hazard response and post-disaster recovery. Thus, for example, in the work of C. Hong et al., an open platform with GIS/BIM integration, real-time monitoring, visualization, and even a VR component for training/experience functions was investigated. This confirms the practical feasibility of a multilayer digital architecture as a whole [20].

Separately, it should be noted that, in the remediation direction, a significant body of scientific research is presented (including that of the authors of this monograph), which is devoted to contaminated sites, brownfields, and the selection of cleanup/recovery technologies, including DSS and MCDA (Multi-Criteria Decision Analysis) approaches [21]. At the same time, reviews on DSS for brownfield/contaminated land management emphasize the importance of integrating disciplines and

supporting complex decisions, but, as a rule, do not proceed to the process orchestration of large-scale post-conflict logistical operations. It is precisely in this that the subject shift in contemporary research consists, namely, the comprehensive transdisciplinary study of the problematics of remediation + logistics + post-war recovery + digital orchestration.

Thus, the problematics of digital management of the elimination of the consequences of emergency situations (ES) and post-crisis recovery have been developing rapidly in recent years; however, predominantly along separate directions: process management, geographic information systems, artificial intelligence, decision support systems, platform-based data integration, and visualization.

Along with this, the authors of the monograph established the fact that much less frequently mentioned in publications is a unified architecture applied (or proposed for use), where:

– pattern recognition forms diagnostic events with confidence/uncertainty;
– while they would be translated into GIS layers;
– then into executable managerial actions through DSS + BPMN/CPM;
– with an execution and feedback loop;
– and with an end-to-end HMI layer of AR/VR + Human-in-the-Loop.

According to the vision of the authors of this study, the hypothesis that precisely such a linkage for remediation logistics (especially in the post-war context) appears reasonably relevant and may be positioned as an architectural-methodological novelty.

In contrast to works considering pattern recognition, GIS platforms, workflow modeling, and AR/VR support predominantly as separate or weakly connected tools, this study proposes an integrated architecture of a closed digital diagnostic-managerial loop of remediation logistics. In it, the results of Pattern Recognition and geospatial diagnostics are formalized in the form of diagnostic events and, through the Decision & Orchestration Core (DSS + BPMN/CPM), are transformed into executable managerial actions with a verification loop and adaptive feedback, under end-to-end support of AR/VR and Human-in-the-Loop mechanisms.

Based on the analysis of the literature, the following research gaps can be identified:

– *the architectural gap*, which is manifested in the absence of a sufficiently clearly formalized concept of a closed digital diagnostic-managerial loop that organically unites surveillance → pattern recognition → geospatial integration → orchestration core → execution monitoring → feedback learning into a unified logical system for remediation logistics;

– *the operational gap*, characterized by the fact that in most studies the mechanism for transforming recognition results (namely: detection/segmentation/classification/change detection) into executable managerial actions is insufficiently disclosed,

taking into account the priorities of remediation tasks, the timeframes of their execution, the availability of necessary resources, and interagency coordination;

– *the interface gap (HMI gap)*. Thus, AR/VR and Human-in-the-Loop mechanisms are rarely described as an end-to-end layer that ensures the interpretation, coordination, and verification of decisions between analytics, management, and field execution;

– *the domain-applied gap*. Thus, it should be recognized that there exists a deficit of publications in which the logic of process orchestration (BPMN/CPM) and digital diagnostics is systemically adapted precisely to post-war remediation logistics, and not only to natural disasters, industrial facilities, or individual contaminated sites.

4.3 Materials and methods of the study

The authors emphasize that the present study has a conceptual-methodological character and is aimed at developing the architecture of a closed digital diagnostic-managerial loop for post-war remediation logistics, rather than at the empirical validation of a specific software prototype.

The methodological basis of the study was formed on the basis of systemic, comprehensive, and integrative approaches. The systemic approach in the monograph was used for the in-depth study of remediation logistics as a multicomponent managerial system, including observation, diagnostics, geospatial interpretation, decision-making, execution, verification, and feedback. In the context of the problematics under study, the comprehensive approach allowed to take into account the interdisciplinary character of the subject field, uniting the tasks of eliminating the consequences of emergency situations, including the post-war recovery of territories (as is the case with Ukraine), transport-logistical support, and digital management. The integrative approach was applied by the authors' collective of this monographic study for the synthesis of heterogeneous technological and methodological components into a unified architectural framework oriented toward the transition from digital diagnostics to executable managerial actions.

As the basic analytical method, an analysis of the degree of study of the problem and a comparative analysis of publications in the following directions were used: surveillance systems, pattern recognition, GIS/DSS, process/workflow orchestration, BPMN, CPM, AR/VR, and Human-in-the-Loop in the context of emergency situations, post-disaster recovery, and remediation. On the basis of the conducted analysis, the identification of research gaps between the existing solutions in the field of digital diagnostics and the tasks of process-managerial implementation of adopted

managerial decisions was carried out. The results of the literature analysis served as the basis for formulating the research hypothesis about the necessity of a closed digital diagnostic-managerial loop as an integrated architecture for the management of remediation logistics.

The key method of concept construction was architectural-conceptual modeling, which was focused on the formalization of the structure and logic of interaction of the functional blocks of the system. At this stage of the study, the authors of the monograph used the functional decomposition of the initial extended structure with the subsequent optimization of interconnected macro-blocks: collection of observation data, recognition and diagnostic analytics, geospatial integration, decision-making and orchestration core, execution monitoring and verification, feedback, and adaptive refinement. The authors made an attempt, with the help of decomposition analysis, to optimize as much as possible and, along with this, to reduce structural redundancy, as well as to increase the conceptual clarity of the architecture without loss of substantive completeness.

For the formalization of the managerial component, the authors' collective used a process-oriented approach relying on Business Process Model and Notation (BPMN) as a tool for structuring, orchestration, and control of the sequence of actions in remediation logistics. BPMN in the study is considered not only as a language of visual description of processes, but also as a basis for building the executable logic of interaction of all participants at different levels of the management system, the algorithm for task execution, and the determination of control points. In addition to this, the critical path method (CPM) was applied for the purpose of conceptualizing the temporal structure of works, dependencies between the performed tasks, critical nodes, and constraints in terms of established deadlines and actually available resources under conditions of the prevailing scarcity of post-war recovery.

Methods of geospatial integration (GIS) and decision support systems (DSS) were also included in the research framework as tools for the operationalization of recognition results in the management space. Within the framework of conceptual modeling, GIS was used for the representation of georeferenced diagnostic events, risk zones, access constraints, as well as the identification of damage layers and routing parameters. DSS was applied by the authors as a mechanism of multicriteria prioritization and selection of managerial alternatives, taking into account resource scarcity, temporal constraints, and the constantly changing situation. Artificial intelligence (AI) technologies, including pattern recognition, are considered in this monograph as a source of diagnostic events and confidence/uncertainty assessments, which are necessary for launching and adapting the managerial loop.

A separate methodological significance is held by the conceptualization of AR/VR and Human-in-the-Loop mechanisms as an end-to-end HMI layer. For this purpose, the monograph used the approach of functional mapping of the roles of AR/VR/HITL across the phases of the loop, namely: interpretation of diagnostics, coordination of decisions, support of execution, verification, and learning (**Fig. 4.1**). As a result, this made it possible to avoid their reduction to exclusively visualization means and instruments. Ultimately, such an approach ensured a more rigorous integration of human-machine interaction into the management architecture and strengthened the substantiation of the scientific novelty proposed in the monograph.

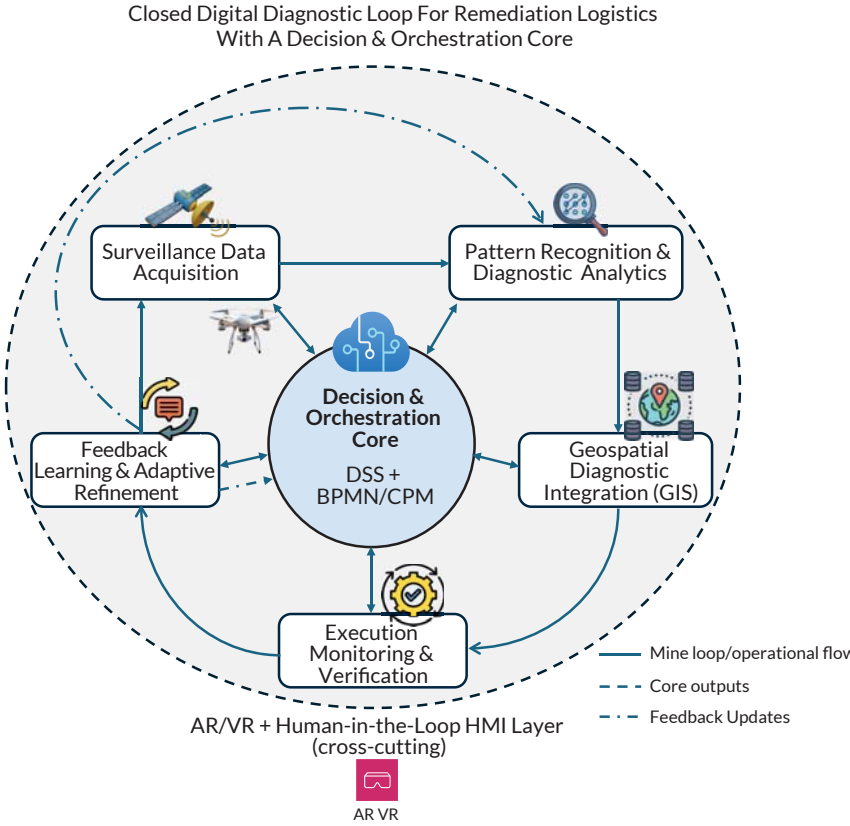


Fig. 4.1 Closed digital diagnostic loop for remediation logistics with a Decision & Orchestration Core

In addition to the above-described methodological techniques, the method of logical synthesis was also applied for the purpose of constructing the transition from diagnostic events to executable managerial actions (the diagnostics-to-action logic) within the framework of the Decision & Orchestration Core (DSS + BPMN/CPM). On this basis, a conceptual model of a closed digital diagnostic-managerial loop for post-war remediation logistics was formed, with the identification of a dual-loop core and domain-specific profiles of further development.

Thus, the totality of the materials and methods used in the monograph provides a sufficient methodological basis for substantiating the proposed architecture and formulating the tasks of its subsequent scenario-based and empirical validation in the authors' further studies.

4.4 Conceptual architecture of the closed digital diagnostic-managerial loop

First of all, it should be noted that this study is a continuation of the in-depth study published earlier, namely on APRe-TISRM, in which the authors proposed [21]:

– the S&D → BIM/DT → 4D/5D + BPMN → execution → post-monitoring loop;
– the role of Surveillance & Diagnostics;
– pattern/image recognition;
– Human-in-the-Loop;
– gates (Confidence/Eco/Safety/Audit);
– the translation of observation/recognition results into manageable actions.

That is, the authors fundamentally laid down: the diagnostic input, the semantic core, the orchestration of processes, as well as the execution and feedback loop. This study makes the next step, namely: it conceptualizes the closed digital diagnostic-managerial loop, strengthens the role of Pattern Recognition as part of the architecture, and introduces AR/VR as a cross-cutting HMI layer. Thus, this is a deepening of the architectural decomposition. It is very important to emphasize that BIM/DT and AR/VR in the logic of the study are not mutually exclusive technologies, but orthogonal and complementary ones. Thus, BIM/DT is responsible for semantic/operational state representation, and AR/VR is responsible for human-centered decision interaction and execution support.

One of the key tasks of the authors is the reflection of a logical scientific evolution:

1. *Stage* 1 *(BIM/DT-oriented remediation):* creation of the semantic-operational architecture of APRe-TISRM, metrics, interoperability, BOR-index, the S&D → → BIM/DT → Ops linkage.

2. *Stage 2 (Pattern recognition → remediation management in closed loop):*

– expansion to a closed diagnostic-managerial loop with the explicit identification of:

a) PR analytics block;

b) geospatial diagnostic integration;

c) Decision & Orchestration Core;

d) execution monitoring;

e) feedback learning;

f) AR/VR + HITL HMI layer (**Fig. 4.2**).

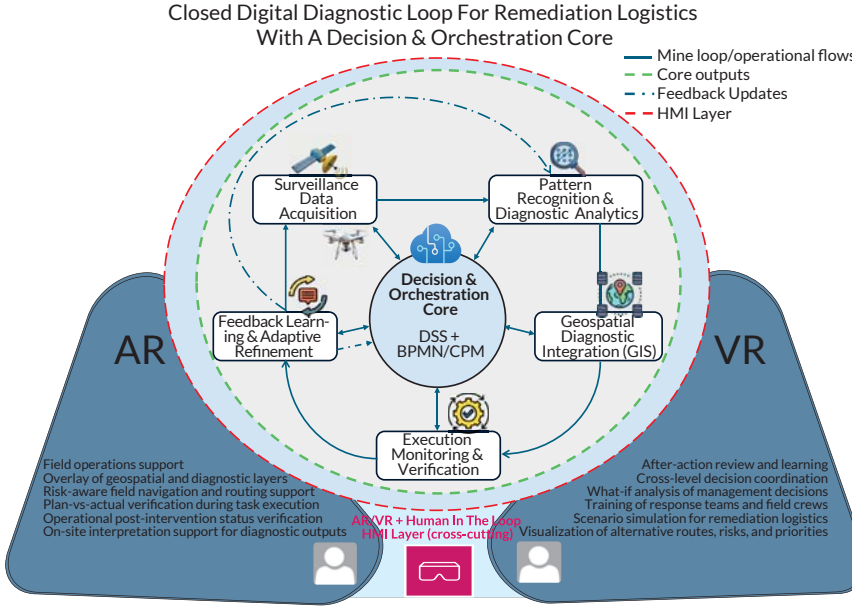


Fig. 4.2 Conceptual model of the closed digital diagnostic-managerial loop for remediation logistics: from pattern recognition to the orchestration and execution of decisions

It is absolutely logical that, as a result, an expansion of the research horizon occurred, a shift of emphasis from the object-semantic and interoperable architecture (BIM/DT) to the human-machine and process-orchestration layer (AR/VR + + HITL + DSS/BPMN/CPM), while preserving BIM/DT as an important element of the integrated architecture.

It is fundamentally important to note that this study does not cancel the BIM/DT-oriented model, but develops it in the direction of adaptive architectural reconfigurability and expands the functional loop through the HMI layer (AR/VR + HITL), ensuring cognitive and operational connectedness.

For greater understanding of the scientific position of the authors, it is proposed a decomposition of the general architecture with the identification of a dual-loop core:

1. *Loop* 1 *(semantic-operational)* is responsible for:
– representation of the state of objects/territories;
– geospatial referencing;
– traceability of state and changes;
– coordination of digital entities and field observations.
2. *Loop* 2 *(decision orchestration)* is responsible for:
– prioritization;
– planning of resources/timeframes;
– formation of executable tasks;
– workflow management and execution control.

It is precisely the dual-loop core that best explains reconfigurability:

– recognition algorithms can be changed without breaking the traceability loop;
– the orchestration logic (BPMN/CPM/DSS) can be changed without breaking the BIM/DT/GIS state model;
– the HMI (AR/VR/HITL) can be strengthened without changing the fundamental data semantics;
– components can be switched on/off depending on the type of emergency situation and the phase of works.

That is, reconfigurability becomes a formally substantiated property. Within the framework of the study, reconfigurability is considered as a systemic property of the loop, ensuring the controlled restructuring of diagnostic, geospatial, orchestration, and HMI components without loss of the integrity of the logic of decision-making and execution.

For the purpose of clarifying the boundaries of applicability of the architecture, it is expedient to distinguish three modes of its configuration:

– BIM/DT-centric mode (for the management of individual infrastructure objects, where the semantic-operational layer dominates);
– loop-centric orchestration mode (for territorial multi-actor remediation logistics, where the central role is played by the DSS + BPMN/CPM core);
– hybrid mode (BIM/DT + loop + AR/VR/HITL) for complex post-crisis and post-war scenarios (**Fig. 4.3**).

Such differentiation confirms that differences in technological emphases reflect not a contradiction of approaches, but a manifestation of the adaptive architectural reconfigurability of a unified diagnostic-managerial loop.

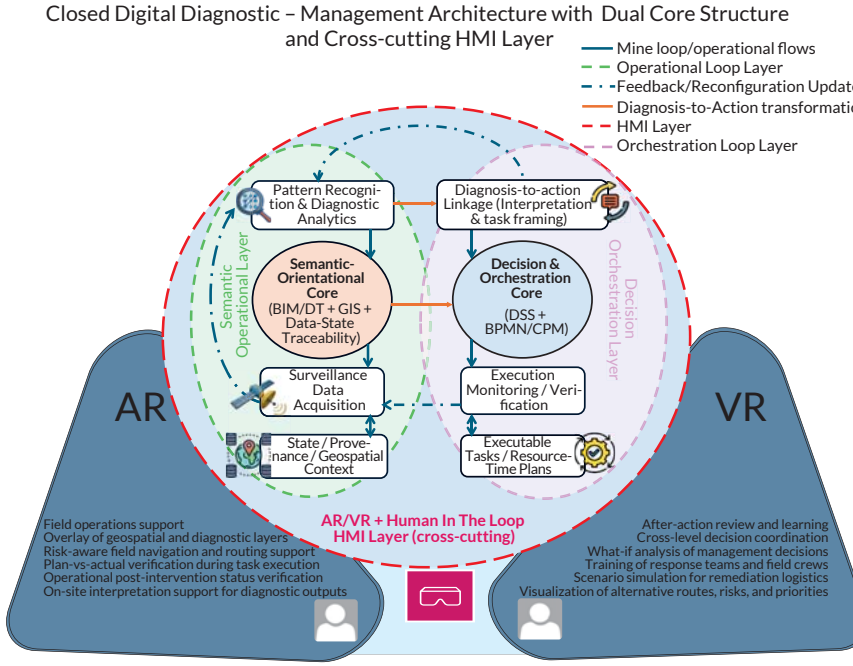


Fig. 4.3 Dual-core architecture of the closed digital diagnostic-managerial loop for remediation logistics: semantic-operational core, orchestration core, and cross-cutting HMI layer

The representation of the extended architecture in the form of a formalized model with a dual-loop nature of the core is logical and methodologically expedient, since it makes it possible to eliminate the contradiction concerning the relationship between BIM/DT, GIS, data-state traceability, and the decision orchestration core (DSS + BPMN/CPM). Such an approach demonstrates that the indicated components do not compete with each other, but form functionally distinguishable, yet interrelated levels of a unified system: the semantic-operational layer ensures state representation, geospatial referencing, and the traceability of changes, whereas the decision orchestration core transforms diagnostic results into executable actions

taking into account timeframes, resources, and workflow constraints. In this architecture, AR/VR + Human-in-the-Loop mechanisms are naturally interpreted as a cross-cutting HMI layer ensuring cognitive and operational connectedness between diagnostics, management, and execution. Formalization in such a formulation also strengthens the substantiation of the property of adaptive architectural reconfigurability, making it possible to describe the change in the configuration of the loop depending on the type of emergency situation, the recovery stage, and the quality of available data.

Summarizing all of the above-presented, it should be noted that the previous and the present studies may be presented as consecutive stages in the development of a unified concept of digital management of the elimination of the consequences of emergency situations and the remediation of territories, where the first work lays the BIM/DT-oriented foundation of reengineering and traceability, and the second one (that is, the present study) expands the architecture to a closed diagnostic-managerial loop with an emphasis on pattern recognition, orchestration, and human-centered execution support.

From a qualitative perspective, the proposed approach differs in that it integrates previously fragmented components (Pattern Recognition, GIS, DSS, BPMN/CPM, as well as AR/VR and Human-in-the-Loop mechanisms) into a unified closed management architecture. The proposed dual-core model, combining the semantic-operational layer and the decision and orchestration core, ensures structural flexibility and allows adaptive reconfiguration of the system without compromising its integrity. At the same time, the cross-cutting HMI layer enhances cognitive, coordination, and operational connectedness between the stages of diagnostics, decision-making, and execution. As a result, the approach enables not only data analysis but also its consistent transformation into executable and verifiable managerial actions, which fundamentally distinguishes the proposed architecture from fragmented digital solutions.

4.5 Metric representation of the closed loop and measurement of the synergy of the dual-loop core

After substantiating the conceptual architecture of the closed digital diagnostic-managerial loop, there arises a necessity to move from structural description to the presentation of its measurable effects and mechanisms of practical implementability. For this purpose, the authors introduce a metric representation of the "diagnosis-to-action-to-verification" cycle and indicators that will make it possible to quantitatively capture the degree of closure of the loop, as well as the synergy of the

dual-loop core through traceability closure. Such an evaluative framework makes it possible to set criteria by which the managerial effectiveness of the integration of the semantic-operational layer and the decision orchestration core may be interpreted, moreover, in different configuration modes.

The authors of the study decided to form a step-by-step metric scheme coordinated with the structure of the loop and its dual-loop core (**Table 4.1**).

At the first stage, the recognition system forms a flow of diagnostic events on the basis of observation data and simultaneously evaluates the degree of their reliability. This makes it possible to create the initial informational basis for making subsequent managerial decisions. At the second stage, diagnostic events are transformed into specific executable tasks, each of which receives resource, temporal, and process formalization and constraint. In other words, a transition takes place from the identification of a problem to the formation of an optimal action plan. At the third stage, control and confirmation of the actual implementation of the tasks set earlier are carried out. However, it should be especially emphasized that, for the full closure of the managerial cycle, it is not sufficient merely to confirm their execution. Along with this, it is also necessary to record the proven change in the state of objects, territories, or affected sites in the semantic-operational environment. It is precisely this stage that completes the full cycle "diagnostics – action – verification – state updating" and makes management truly closed.

It is fundamentally important that such a result is possible only through the joint operation of the two parts of the architecture. In the proposed configuration, the first is responsible for translating diagnostic data into managerial actions, while the second is responsible for the recording, updating, and traceability of the state after the execution of the complex of remediation works. Therefore, synergy is manifested not simply in the execution of a set of tasks differing in character, but in the ability of the system to bring diagnostic events to a confirmed and traceable result. In this context, it should be emphasized that, in order to assess the effect of such integration, it will be expedient to compare different modes of operation of the architecture, namely: with the predominance of the semantic-operational loop, with the predominance of the decision orchestration loop, as well as in a hybrid configuration. The synergistic effect will be considered achieved only in the case when the joint operation of the two loops ensures a more complete and evidentiary completion of the managerial cycle, in comparison with the operation of each of them separately. This approach makes it possible to connect the conceptual architecture directly with the practical assessment of its operation and creates a basis for further empirical and simulation-based verification in the tasks of remediation and post-crisis recovery (and not only in the transport-logistics sphere).

Table 4.1 Step-by-step metric representation of the closed loop and measurement of the synergy of the dual-loop core

No.	Loop step (structural)	Formalized object/set	Metric/calculated Indicator	Functional meaning/application
1	2	3	4	5
1	PR generates diagnostic events and uncertainty	D_T, (N_{diag}=	Let R_T – be the initial flow of "raw" diagnostic triggers from Pattern Recognition/Digital Diagnostics over period T (detections, segmentations, change detection, etc.). Each trigger $r \in R_T$ has a minimal set of attributes $r = <t, g, c, p, u, s>$, where t – time (timestamp); g – geolocation (specific place/polygon/object); c – class/type of event (damage/contamination/blockage/change, etc.); $p \in [0,1]$ – confidence; $u \in [0,1]$ – uncertainty; $s \geq 0$ – severity/impact score (importance/severity; may be from model, DSS, or rules). It is expedient to introduce a rule of "managerial significance" $D_T = \{r \in R_T \mid p(r) \geq \tau_p \wedge u(r) \leq \tau_u \wedge s(r) \geq \tau_s\}$, where τ_p, τ_u, τ_s – thresholds (may be fixed or reconfigurable by modes)	Pattern Recognition/Digital Diagnostics forms events d_i with attributes: t, g, type, conf/unc
2	Core transforms events into tasks	A_T, (N_{act}=	Let A_T – be the set of executable remediation-workflow tasks created by the decision orchestration core (DSS + BPMN/CPM) over period T as a managerial response to diagnostic events. Each task $a \in A_T$ is described by the minimal attribute set $a = <id, t_c, g, \theta, \pi, k, \rho, \sigma>$, where id – task identifier; t_c – creation timestamp; g – geolocation/object (site/asset); θ – task type (inspection/clearance/repair/remediation/logistics delivery/etc.); π – priority; k – criticality/risk; ρ – resource requirements (skills/equipment/materials); σ – status (created/assigned/in-progress/completed/verified/rejected). Optimized metric N_{act}: $N_{act} = \left\lvert A_T^{diag} \right\rvert$, $A_T^{diag} = \left\{ a \in A_T \mid \exists d \in D_T^* : a \in f(d) \right\}$	Decision & Orchestration Core (DSS + BPMN/CPM) applies prioritization rules as well as orchestration and creates workflow tasks

Continuation of table 4.1

1	2	3	4	5
3	Execution monitoring confirms completion	V_T, (N_{ver}=	Verified tasks (confirmed execution) $V_T = \{a \in A_T^{diag} \mid veryfy(a) = 1\}$. Number of verified tasks $N_{ver} = \lvert V_T \rvert$. Verification ratio $VR = \frac{N_{ver}}{N_{act}}$ or $VR = \frac{\lvert V_T \rvert}{\lvert A_T^{diag} \rvert}$	Execution Monitoring/Verification records the fact of the implementation of remediation tasks and their compliance with requirements
4	SO-layer records state updates	T_T, (N_{tr}=	Traceability state updates (as-is → as-done/as-remediated) $T_T = \{a \in V_T \mid trace\_update(a) = 1\}$. Number of confirmed traceability updates $N_{tr} = \lvert T_T \rvert$	Semantic-Operational layer (BIM/DT + GIS + traceability) records the update "as-is → → as-done/as-remediated" (evidentiary traceability)
5	Assessment of closure quality via traceability	TR	Traceability ratio $TR = \frac{N_{tr}}{N_{ver}}$	Shows maturity of integration: how many verified works result in evidentiary state updates (audit, reporting, quality management)
6	Interpretation of dual-core synergy	–	Metric representation of synergy via $D2A_I^{(tr)}$ and TR: $D2A_I^{(tr)} = \frac{N_{tr}}{N_{diag}} = \frac{\lvert T_T \rvert}{\lvert D_T \rvert}$ and $TR = \frac{N_{tr}}{N_{ver}} = \frac{\lvert T_T \rvert}{\lvert V_T \rvert}$	The indicators capture the share of diagnostic events that ended with an evidentiary state update (traceability), and the share of verified works anchored in the semantic-operational loop. Therefore, it is precisely they that are the most direct metrics of the synergy of the dual-loop core
7	Mode-based comparison (SO/DO/Hybrid)	$D2A_I^{(tr)}(SO)$, $D2A_I^{(tr)}(DO)$, $D2A_I^{(tr)}(Hybrid)$	Synergistic increment $\Delta_{syn} = D2A_I^{(tr)}(Hybrid) - \max\{D2A_I^{(tr)}(SO), D2A_I^{(tr)}(DO)\}$	Shows measurable "added" integration effect: hybrid mode provides more complete traceability closure compared to dominance of a single loop

From a quantitative perspective, the proposed architecture introduces a metric framework that enables the evaluation of not individual analytical stages, but the full cycle "diagnosis – decision – execution – verification – state update". In particular, it allows measuring the proportion of diagnostic events transformed into executable managerial actions, successfully verified, and completed with evidence-based and traceable state updates. This makes it possible to quantitatively identify gaps between analytics, orchestration, and execution, as well as to provide a basis for comparing different architectural configurations and evaluating the synergistic effect of their integration.

4.6 Expansion of the role of AR/VR and Human-in-the-Loop mechanisms in the architecture of the closed digital diagnostic-managerial loop

Within the framework of the proposed concept of a closed digital diagnostic-managerial loop, it is expedient to reconsider the traditional role of AR/VR technologies and Human-in-the-Loop (HITL) mechanisms, shifting the emphasis from their auxiliary visualization function, which is traditionally highlighted in the scientific literature, to a systemically significant level of the human-machine interface (HMI-layer). Thus, in the predominant majority of applied solutions, AR/VR are used mainly as means of representing the actual situation in cases of personnel training or demonstration of scenario development; however, under conditions of eliminating the consequences of emergency situations and implementing remediation measures, such an approach is insufficiently studied and covered. This is due to the fact that the key problem of modern practice lies not only in the lack of data, but also in the gap between digital diagnostics, managerial interpretation, and the execution of planned decisions in field conditions. In this regard, AR/VR and HITL, according to the authors of the monograph, should be considered as a cross-cutting layer that will ensure the cognitive connectedness of the loop "observation – recognition – geospatial diagnostics – decision – execution – verification – adaptation". The expansion of the functional role horizon as a whole is driven by the necessity of reducing the risks of erroneous interpretation of recognition results, increasing the consistency of multi-level management and coordination, as well as ensuring a verifiable transition from diagnostic events to effectively managed actions. Thus, AR/VR and HITL acquire the status not of peripheral technologies, but of an architecturally embedded mechanism for increasing the reliability and adaptability of the management of the elimination of the consequences of emergency events.

At the phase of surveillance and pattern recognition, the cross-cutting HMI layer provides interpretative support for the results of intelligent diagnostics, including visual identification of detected damage to objects, risk zones, as well as the scale of required interventions, and also the recording of changes in the state of the remediation object taking into account confidence/uncertainty assessments. At this phase, HITL mechanisms perform the function of expert validation of critically significant or low-reliability diagnostic events, thereby making it possible to minimize the probability of the transition of false-positive or false-negative results into the managerial loop. At the phase of geospatial integration (GIS/DSS), AR/VR and HITL support the coordination of the spatial context of decisions, namely: visual analysis of access constraints, priority intervention zones, as well as alternative routes for the delivery of various types of resources and the scenario consequences of a specific choice. In this context, VR acquires particular importance as an environment for scenario modeling and collective examination of alternative decisions in the what-if analysis mode, as well as a tool for interagency cooperation and coordination of efforts at the phase of project analysis and before the beginning of field implementation of remediation/reconstruction/recovery works. At the phase of the Decision & Orchestration Core (DSS + BPMN/CPM), HITL mechanisms ensure the controlled inclusion of an expert in the process of making specific decisions where confirmation of the level of safety, prioritization, change of the sequence of remediation tasks, redistribution of resources, or approval of a modification scheme of the process is required. It should be especially emphasized that, at this phase, AR/VR do not replace the orchestration core, but ensure its cognitive transparency, while increasing the quality of interpretation of the consequences of managerial scenarios.

At the phase of Execution Monitoring & Verification, AR acquires the greatest practical significance as a tool of field support in situations of: overlaying risk layers on terrain, routing the delivery of resources and equipment, indicating the statuses of task execution, territorial analysis of the boundaries of intervention zones, as well as analysis of infrastructural constraints. This makes it possible to reduce the level of errors in execution, increase the compliance of actual actions with the previously formed plan, and accelerate the confirmation of completed operations. HITL at this phase makes it possible to verify the obtained results, record deviations, and, if necessary, make a decision on re-inspection and/or corrective actions. At the phase of Feedback Learning & Adaptive Refinement, VR and HITL may be used for post-scenario analysis, as well as for the purpose of modifying the course of remediation operations, identifying the causes of delays, routing conflicts, and interpretation errors, as well as for adapting procedures, confidence thresholds, and BPMN regulations. Thus, AR/VR and HITL become not only means of supporting the entire

cycle of implementation of remediation measures, but also tools of management, organizational learning, and a base of practical knowledge within the management loop. Taken together, all of the above forms continuous human-machine support at all stages of the lifecycle of remediation logistics.

Summarizing the description of this approach, it may be stated that the scientific novelty lies in the conceptualization of AR/VR and Human-in-the-Loop mechanisms as a cross-cutting HMI layer in the architecture of the closed digital diagnostic-managerial loop, rather than as local tools of visualization or training. The proposed expansion of the role makes it possible to formally connect digital diagnostics, process orchestration, and the field implementation of the complex of remediation measures through a unified loop of interpretation, coordination, and verification of decisions. In contrast to fragmented approaches, where AR/VR and expert participation are integrated in most cases episodically, in this concept they act as a systemic mechanism for increasing the resilience of management under conditions of uncertainty, time scarcity, and high dynamics of situational changes. This strengthens the architectural integrity of the model proposed by the authors of the monograph and substantiates its applicability to the tasks of eliminating the consequences of emergency situations of a natural, technogenic, and military nature, as well as to post-crisis remediation measures (**Table 4.2**).

This matrix appears methodologically strong and useful, as it performs not merely an illustrative, but a substantiating function within the structure of the study. First, it demonstrates the functional continuity of human-machine interaction throughout the entire loop, namely: from the interpretation of observation results and image recognition to the verification of damage, from training to the modification of the remediation logistics route. This means that human participation and immersive interfaces are embedded not in a single separate stage, but in the logic of the entire management lifecycle. Second, the matrix confirms the differentiation of the roles of AR, VR, and HITL. That is, the matrix demonstrates that the technologies are distributed not arbitrarily, but in accordance with their most rational function within the management architecture. Third, the table confirms that the concept proposed by the authors indeed ensures cognitive, coordination, and operational connectedness between digital diagnostics, decision-making, process orchestration, as well as the field implementation of the set of remediation/reconstruction/recovery measures, verification of results, and subsequent adaptive training of personnel. This is particularly important, because it is precisely such connectedness that is usually lacking in many existing digital solutions, where analytics, management, and execution remain disconnected. Fourth, the matrix formed by the authors confirms the thesis that the HMI layer is not an auxiliary, but an architecturally significant mechanism for increasing the resilience of crisis management under conditions of uncertainty, time scarcity, high dynamics of

situational change, increasing the effectiveness of interagency cooperation and coordination of joint efforts for the purpose of the prompt elimination of the consequences of emergency events, as well as the necessity of a rapid transition from diagnostics to real actions. Fifth, this matrix also indirectly confirms the closed nature of the loop itself, since the functions of AR/VR and HITL are distributed not only at the input and in the middle of the process, but also at the final stages, namely at the stages of verification, learning, and adaptive refinement. And this means that the loop is indeed considered as a cycle, rather than as a linear chain of actions.

Table 4.2 Matrix of functional distribution of AR/VR and Human-in-the-Loop mechanisms by phases of the closed digital diagnostic-managerial loop of remediation logistics

Loop phases/ HMI functions	Interpretation	Validation	Coordination	Scenario analysis	Field support	Verification	Learning/ adaptation
Surveillance Data Acquisition	± (coverage control)	–	–	–	–	–	–
Pattern Recognition & Diagnostic Analytics	AR/VR + HITL	HITL	–	–	–	–	–
Geospatial Diagnostic Integration (GIS)	AR/VR	HITL	AR/VR + HITL	VR	–	–	–
Decision & Orchestration Core (DSS + + BPMN/ CPM)	–	HITL	AR/VR + HITL	VR + HITL	–	–	–
Execution Monitoring & Verification	AR	HITL	–	–	AR + HITL	AR + HITL	–
Feedback Learning & Adaptive Refinement	VR + HITL	HITL	HITL	VR	–	HITL	VR + HITL

In general form, presented below is an aggregated list of what is included in AR/VR as parts of the cross-cutting HMI layer:

1. Augmented Reality for field operations support, in particular for: overlaying digital layers on terrain/a tablet screen (including, among other things,

the designation of risk zones, routes, statuses of tasks being performed, as well as access restrictions).

2. Augmented Reality for geospatial data visualization, including for the display of GIS layers and diagnostic objects referenced to real space.

3. Augmented Reality for plan-vs-actual verification, in the context of comparing planned remediation tasks with the work actually performed on site.

4. Augmented Reality for field navigation and routing support, including visual prompts regarding safe routes, access points, as well as bypassing hazardous zones.

5. Virtual Reality for scenario simulation, including the playing out of alternative scenarios of work execution, logistical routes, and sequences of actions.

6. Virtual Reality for team training and preparedness, which may be successfully used for practicing response mechanisms, coordination, and task execution under conditions as close as possible to real ones.

7. Virtual Reality for what-if analysis of management decisions, which may be used for visual verification of the consequences of changing priorities, timeframes for the implementation of remediation measures, routes for resource delivery, and their distribution.

8. Virtual Reality for cross-level decision coordination, which may be applicable in training of joint scenarios for the strategic, regional, and field levels of management of different types of remediation/reconstruction/recovery measures.

9. Immersive visualization of diagnostic outputs, which consists in the presentation of the results of recognition/segmentation/detection in a form maximally understandable for operational human interpretation.

10. Immersive visualization of model uncertainty and confidence, including the display of confidence/uncertainty estimates for reducing the risk of erroneous interpretation of AI analytics.

11. AR/VR as a human-machine interface layer in the loop. Let's emphasize that this is not merely visualization, but a means of interpretation, coordination, as well as confirmation and verification of decisions.

12. AR/VR as a means of after-action review and learning. This refers to the application of VR/AR in the feedback phase for the analysis of errors, refinement of processes, and team training [22, 23].

Thus, it is possible to make a generalization: AR in the proposed architecture is predominantly oriented toward supporting field execution and verification, whereas VR is oriented toward scenario planning, decision coordination, and learning (AR supports execution and verification; VR supports simulation, coordination, and learning).

As a result, there is every basis to present a kind of "Evolutionary Ladder" of the development and expansion of the role of AR/VR:

Visualization tool → Situational awareness support → Decision support interface → Execution support interface → Verification interface → Cross-cutting HMI layer in closed-loop management (**Fig. 4.4**).

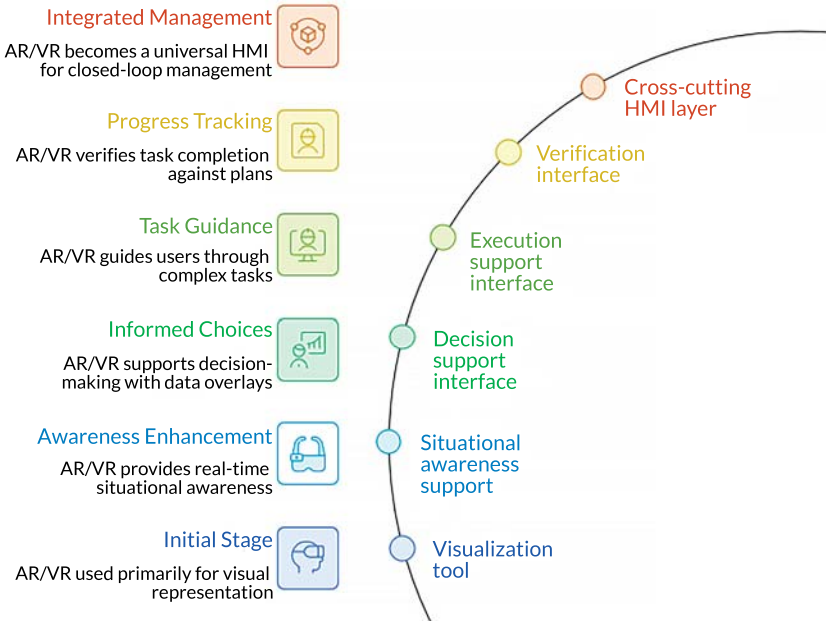


Fig. 4.4 Evolutionary ladder of the AR/VR role in remediation logistics management

4.7 Conclusions

The authors have proposed a conceptual architecture of a closed digital diagnostic-managerial loop for remediation logistics, in which the Pattern Recognition block is formalized as a source of diagnostic events and uncertainty, integrated into the geospatial layer and further transformed into executable managerial actions through the Decision & Orchestration Core (DSS + BPMN/CPM), with cross-cutting support of the HMI layer (AR/VR + Human-in-the-Loop). One of the key properties of the proposed architecture is its adaptive architectural reconfigurability,

understood as the ability of the closed digital diagnostic-managerial loop to change the configuration of functional blocks, the connections between them, and the parameters of process orchestration depending on the type of emergency situation, the stage of remediation, and resource constraints. The proposed approach demonstrates both quantitative and qualitative advantages compared to fragmented solutions. From a quantitative perspective, it enables the evaluation of the degree of closure of the full management cycle from diagnostic events to evidence-based and traceable state updates. From a qualitative perspective, it ensures the integration of previously fragmented components into a unified adaptive architecture capable of transforming diagnostic results into executable and verifiable managerial actions.

The monograph presents an analysis of the expansion of the horizon of applicability of the proposed conceptual approach to the tasks of eliminating the consequences of emergency situations of a natural, technogenic, and military nature, while preserving the transport-logistics sphere as the priority domain of implementation. The study focused attention on the transport-logistics sphere as the priority contour for approbation and development, since it is precisely in it that the necessity of rapid transformation of diagnostic data into executable managerial actions is most clearly manifested [24].

In addition, the scientific novelty of the study lies in the conceptual rethinking of the AR/VR role in systems for eliminating the consequences of emergency situations and remediation of affected territories: AR/VR are proposed to be considered not as isolated tools of visualization or training, but as a cross-cutting layer of human-machine interaction (cross-cutting HMI layer) in the architecture of the closed digital diagnostic-managerial loop. In contrast to traditional approaches, AR/VR are embedded into a unified loop together with pattern recognition, GIS, DSS, and BPMN/CPM, ensuring the transition from the results of digital diagnostics to executable managerial actions in the tasks of eliminating destruction, reconstruction, construction, repair, and full remediation of territories. It is substantiated that under modern conditions of high uncertainty, scarcity of time and resources, AR/VR acquire new functional characteristics, namely: support for the interpretation of diagnostics, coordination of decisions, support of execution, verification of results, as well as training and adaptive adjustment of the loop. Thus, the range of their application is significantly expanded – from local visual support to an architecturally significant component of the engineering of managerial processes in a multi-actor crisis environment. The proposed approach forms a basis for further formalization and empirical approbation of integrated diagnostic-managerial systems of a new generation in subsequent studies.

Conflict of interest

The authors declare the absence of any conflicts of interest in relation to this monograph chapter and its published results, including financial interests, non-financial personal relationships, as well as other institutional interests that could have influenced the work performed.

Financing

The study was performed without targeted financial support (without specific grants, sponsorship, or commercial funding).

Data availability

This chapter has a conceptual-methodological character and does not rely on empirical data. The materials supporting the results of this study (including conceptual diagrams, metric definitions, and formalized hypotheses, etc.) were developed by the authors and may be provided upon reasonable request.

Use of artificial intelligence statement

The authors used artificial intelligence technologies exclusively for language and academic-style support, specifically for verifying the correctness, consistency, and clarity of the English wording, as well as for minor stylistic refinement. In particular, the authors used the generative language model ChatGPT (OpenAI, GPT-5.2 Thinking) for language checking and limited editorial editing. The authors bear full responsibility for the content of the final version of the manuscript; the AI tool is not credited as an author and is not responsible for the reported results.

Acknowledgments

The authors are grateful to the reviewers for the work performed and to the publisher for the opportunity to publish this scientific study.

Authors' contributions

Tetiana Cherniavska: Conceptualization; Theoretical framework.
Bohdan Cherniavskyi: Methodology; Architectural structuring support.
Rusnak Alla: Interpretation of results; Development of the closed-loop architecture.
Nadtochii Iryna: Writing – original draft; Writing – review & editing.
Artur Harahulia: Visualization (figures and scheme logic).
Oleksii Bobrovskyi: Formal analysis; Analysis of scientific sources.
All authors contributed to the scientific content of this chapter, jointly developed the conceptual idea and structure of the work, and approved the final version of the manuscript.

References

1. Pantiris, P., Pallis, P. L., Chountalas, P. T., Dasaklis, T. K. (2025). Enhancing Coordination and Decision Making in Humanitarian Logistics Through Artificial Intelligence: A Grounded Theory Approach. Logistics, 9 (3), 113. https://doi.org/10.3390/logistics9030113
2. Patnala, P. K., Regehr, J. D., Mehran, B., Regoui, C. (2024). Resilience for freight transportation systems to disruptive events: a review of concepts and metrics. Canadian Journal of Civil Engineering, 51 (3), 237–263. https://doi.org/10.1139/cjce-2023-0187
3. Köstepen, Z. N., Selim, H. (2025). Literature Review on Humanitarian Logistics in Disaster Management: A Risk-Oriented Approach. Sustainability, 17 (21), 9773. https://doi.org/10.3390/su17219773
4. Kong, J., Zhang, C., Simonovic, S. P. (2023). Resilience and risk-based restoration strategies for critical infrastructure under uncertain disaster scenarios. Sustainable Cities and Society, 92, 104510. https://doi.org/10.1016/j.scs.2023.104510
5. Plotnikov, O.; Carrillo-Pina, J., Sharov, O. (Eds.) (2025). The Influence of the Fragmentation of the World Economy on the Post-War Reconstruction of Ukraine. The Geoeconomics of the International Monetary Order. Cham: Palgrave Macmillan, 71–91. https://doi.org/10.1007/978-3-031-90851-4_3
6. Becker, T., Gorodnichenko, Y., Weder di Mauro, B. (2025). How to Rebuild Ukraine: A Synthesis and Critical Review of Policy Proposals. Annual Review of Economics, 17 (1), 293–320. https://doi.org/10.1146/annurev-economics-081324-093012

7. Nivievskyi, O., Goriunov, D., Nagurney, A., Dvornichenko, D., Kopytsia, I., Kirchner, R. et al. (2025). Recovery, Resilience and Resources. Ukrainian Analytical Digest, 14. https://doi.org/10.3929/ethz-c-000784202
8. Cherniavskyi, B. (2024). Challenges of Successful Remediation in Ukraine after the End of Military Activities in the Context of European Integration. Economic development and policies: realities and prospects. European integration, convergence and cohesion. Sofia: ERI at BAS, 103–107.
9. Petrukha, N., Fedirko, N., Piatnychuk, I., Lyashenko, P., Plakhotnii, D. (2025). Economic Rebuilding Frameworks in Post-War States: Takeaways for Ukraine. International Journal of Economic Sciences, 14 (1), 196–210. https://doi.org/10.31181/ijes1412025196
10. Chukwuebuka Ahuchogu, M., Jawad, A. B., Hamidi, I. A., Jayasundar, S., Howard, E. (2025). Real-Time Image-Based Data Processing and its Applications in Managerial Decision-Making and Risk Analysis. Eksplorium, 46 (1), 1552–1565. https://doi.org/10.52783/eksplorium.181
11. Mondal, G., Dhanaraj, R. K., Banerjee, C.; Bhattacharya, A., Dutta, S., Yang, X. S., Bose, S. (Eds.) (2026). The Future of Risk Detection: Integrating RS, GIS, and AI for Holistic Hazard Awareness. Proceedings of International Conference on Computational Intelligence and Information Retrieval. Cham: Springer, 507–522. https://doi.org/10.1007/978-3-032-02790-0_36
12. Cimini, C., Lagorio, A., Cavalieri, S., Riedel, O., Pereira, C. E., Wang, J. (2022). Human-technology integration in smart manufacturing and logistics: current trends and future research directions. Computers & Industrial Engineering, 169, 108261. https://doi.org/10.1016/j.cie.2022.108261
13. Moazzeni, S., Sgarbossa, F. (2025). Collaborative Logistics and Digital Technologies in Rural Contexts: A Systematic Review and a Decision Aid Model for Logistics Decision-Makers. SSRN. https://doi.org/10.2139/ssrn.5236703
14. Rahman, M. A. (2025). Review of Applied Science and Technology. SSRN. https://doi.org/10.2139/ssrn.5360313
15. Rahnemoonfar, M., Chowdhury, T., Murphy, R. (2023). RescueNet: A High Resolution UAV Semantic Segmentation Dataset for Natural Disaster Damage Assessment. Scientific Data, 10 (1). https://doi.org/10.1038/s41597-023-02799-4
16. Albahri, A. S., Khaleel, Y. L., Habeeb, M. A., Ismael, R. D., Hameed, Q. A., Deveci, M. et al. (2024). A systematic review of trustworthy artificial intelligence applications in natural disasters. Computers and Electrical Engineering, 118, 109409. https://doi.org/10.1016/j.compeleceng.2024.109409
17. Rodrigues da Silva, A., Estima, J., Marques, J., Gamito, I., Serra, A., Moura, L. et al. (2023). A Web GIS Platform to Modeling, Simulate and Analyze Flood Events:

The RiverCure Portal. ISPRS International Journal of Geo-Information, 12 (7), 268. https://doi.org/10.3390/ijgi12070268
18. Betke, H., Seifert, M. (2017). BPMN for Disaster Response Processes – A methodical extension. INFORMATIK 2017, 1311–1324. Available at: https://www.researchgate.net/publication/320556286_BPMN_for_Disaster_Response_Processes-A_methodical_extension
19. Kramer, S. W., Jenkins, J. L. (2006). Understanding the basics of CPM calculations: what is scheduling software really telling you? PMI® Global Congress 2006-North America. Madrid: Project Management Institute. Available at: https://www.pmi.org/learning/library/basics-cpm-scheduling-software-axon-8170
20. Hong, C., Park, S., Ju, K., Lee, J. (2024). An Open Disaster Information Platform, Methodology, and Visualization for High-Rise and Complex Facilities. Buildings, 14 (12), 4047. https://doi.org/10.3390/buildings14124047
21. Motamedi, A., Vaudou, S., Leygonie, R., Forgues, D. (2019). Process re-engineering in owner organizations to improve BIM-based project delivery using requirements management platform. 4th International Conference on Civil and Building Engineering Informatics (ICCBEI). Sendai. Available at: https://www.researchgate.net/publication/339630638_Process_Re-Engineering_in_Owner_Organizations_to_Improve_BIM-Based_Project_Delivery_Using_Requirements_Management_Platform
22. Khanal, S., Medasetti, U. S., Mashal, M., Savage, B., Khadka, R. (2022). Virtual and Augmented Reality in the Disaster Management Technology: A Literature Review of the Past 11 years. Frontiers in Virtual Reality, 3. https://doi.org/10.3389/frvir.2022.843195
23. Surtiari, G. A. K., Dalimunthe, S. A., Reksa, A. F. A., Pelupessy, D., Prasojo, A. P. S., Jibiki, Y. et al. (2024). Making Virtual Reality (VR)/Augmented Reality (AR) Possible to Strengthen Disaster Risk Reduction among Communities at Risk of Tsunami. International Journal of Disaster Management, 6 (3), 291–312. https://doi.org/10.24815/ijdm.v6i2.34523
24. Cherniavskyi, B., Blakyta, H., Susidenko, V., Andreichenko, A., Remyha, Y., Podmazko, O.; Cherniavska, T. (Ed.) (2025). Innovative technologies and digital models in the post-war recovery of the transport and logistics system of Ukraine. Economy in the Era of Digital Transformation: Trends, Opportunities and Perspectives. Tallinn: Scientific Route OÜ, 110–143. https://doi.org/10.21303/978-9908-9706-0-8.ch5

CHAPTER 5

Detection of small and camouflaged objects

Dmytro Krytskyi
Elvira Kaidan
Artem Chekhovsky

Abstract

Detecting small and camouflaged objects in aerial images is challenging, especially when data is acquired from drones. In such images, targets often occupy only a very small portion of the frame and can visually blend in with the surrounding terrain. Image quality is also affected by flight altitude, unstable lighting, motion blur, background noise, and intentional camouflage. Therefore, analyzing such data is complex and time-consuming, significantly increasing the likelihood of missing a target object. Therefore, automatic detection is a highly relevant approach to identifying these objects.

This chapter examines the use of deep learning methods for detecting small and camouflaged objects in aerial photos and videos. The study focuses on YOLO-based detectors, as these models combine good detection quality with high processing speed and are suitable for practical applications. Particular attention is given to the comparison of YOLOv8 and YOLOv11. An experimental study was conducted on an annotated dataset created from publicly available video footage. Two model configurations were trained and evaluated. Image resizing by direct stretching was replaced by adding padding to preserve the proportions of objects within the frame. Additionally, the class structure was simplified, reducing ambiguity during training and increasing classification confidence. These changes were tested alongside the transition from YOLOv8n to YOLOv11s.

The results showed that the improved approach provided more stable detection in complex data and significantly reduced training time. The YOLOv11 model demonstrated the best practical results when working with small targets in complex background conditions. The obtained results confirm that modern architectures based on YOLO family models can be effectively applied to automated data analysis

 DOI: 10.21303/978-9908-845-05-0.ch5

and can serve as a foundation for the further development of intelligent decision support systems.

Keywords

Small object detection, camouflaged object detection, aerial imagery, computer vision, deep learning, object detection, YOLOv8, YOLOv11, image preprocessing, padding, dataset annotation, automated recognition systems.

5.1 Introduction

Unmanned aerial vehicles (UAVs) have significantly expanded the capabilities of remote sensing and the rapid collection of geospatial data. Today, they are widely used for environmental monitoring, infrastructure inspection, and military reconnaissance. Their use allows for the acquisition of detailed visual information over large areas in a relatively short time. At the same time, the volume of photo and video data collected by UAV platforms has increased so much that manual analysis is often insufficient, especially in situations requiring rapid decision-making [1, 2].

Detecting small and intentionally camouflaged objects is particularly challenging [3, 4]. Such targets may occupy only a small number of pixels in an image and are difficult to distinguish from the surrounding background. The situation is further complicated by camouflage, variable lighting, weather conditions, and the visual complexity of the terrain itself [3, 5]. As a result, the operator must process large volumes of visually complex data, which increases the likelihood of missing desired objects. In such cases, approaches designed specifically for small-object analysis, including hyper-inference and image slicing, become particularly relevant [4].

These limitations make automated image analysis an important research and practical challenge. In recent years, deep learning methods have demonstrated good results in object detection tasks, especially in scenes where traditional image processing methods are not robust enough. Convolutional neural networks (CNNs) are particularly effective because they can learn informative visual features directly from data rather than relying on hand-crafted descriptors [6]. For the use of such systems onboard UAVs, it is important that such models be able to run in real-time or near-real-time on onboard systems with limited computing resources [7].

Among modern object detection approaches, the YOLO family remains one of the most widely used solutions. Its single-stage detection scheme allows the model to predict object classes in a single, straightforward pass, providing a good balance between speed and accuracy [8]. Because of this, YOLO-based detectors are widely used in video analytics and other applications requiring fast processing [9].

At the same time, model quality depends not only on the detector architecture itself. The final result is also influenced by the composition of the training dataset, the data augmentation strategy, preprocessing methods, and the choice of training parameters. Even relatively small changes in preprocessing can impact how well the system handles noisy data and how reliably it detects objects close to the sensor's resolution limit [10].

The aim of this study is to explore approaches to automatically detecting small and partially camouflaged objects in photos and videos using modern deep learning methods. Particular attention is paid to the performance of YOLO-based models under challenging observation conditions.

5.2 Problem of detecting small and camouflaged objects in aerial imagery

Detecting small and camouflaged targets in aerial imagery data streams is one of the most critical and complex fundamental tasks in computer vision [11]. The specific nature of obtaining visual information from unmanned aerial vehicles (UAVs) gives rise to a number of disruptive factors that significantly reduce the effectiveness of classical segmentation and identification algorithms.

The main challenge in such situations is the gap between the sensor's resolution and the small size of the target object. When capturing a large area, the UAV is flying at a high height, and the objects occupy a very small number of pixels in the frame [12]. As a result, we get a "deficiency": texture, shadows – all the most important criteria for recognition are lost, and the system is unable to distinguish the object from the background.

Camouflage must also be considered. During military or reconnaissance missions, objects are often concealed. During such operations, camouflage nets or patterns are used, or the specific terrain and natural factors are taken into account. In such situations, the spectral characteristics and the surrounding surface become similar. The texture of the target to be identified and the surroundings become so similar that the visual boundary becomes blurred. In such cases, classical gradient detection methods fail to provide reliable results for the model.

An additional challenge is posed by the structural non-uniformity of the background, namely its high entropy. Aerial imagery typically contains a high number of non-textural elements, such as buildings, roads, and so on. These objects create visual noise, leading to the detection of false contours and textures. For a system operating in real time and required to make rapid decisions, such errors are unacceptable.

What also cannot be ignored is that the mission is carried out under dynamic conditions [13]. During operation, the UAV performs a series of maneuvers, resulting in sudden changes in camera angles, lighting fluctuations, and additional atmospheric effects. All of these factors lead to blurring and geometric distortions in the video.

Conventional computer vision methods, which rely on manually created features, prove to be insufficiently robust against the difficulties mentioned earlier. In comparison, deep learning allows for the automatic construction of a feature hierarchy. It gradually adapts to complex spatial structures and nonlinear dependencies in the data.

Given all the aforementioned problems, the task is to create adaptive neural network approaches based on deep learning, tailored to the specifics of aerial images from UAVs. The system must ensure high accuracy with a low false positive rate and robustness to dynamic imaging conditions.

5.3 Deep learning approaches for small object detection

Serious changes have taken place in the field of computer vision over the past ten years. Instead of deterministic algorithms, in which features were specified manually, deep learning methods have increasingly come into use.

The fundamental mechanism for feature extraction in such systems is the convolution operation, which preserves spatial correlations between pixels.

The convolution operation can be formally represented as follows

$$(I*K)(i,j)=\sum_m\sum_n I(i+m,j+n)K(m,n), \tag{5.1}$$

where I – the input image; K – the convolution kernel; i, j – the pixel coordinates in the output feature map.

As data passes sequentially through a cascade of convolutional layers, low-level features (lines, gradients) are transformed into high-level semantic structures.

Activation functions are used to model nonlinear dependencies and improve the network's approximation capability. The standard for modern architectures is the Rectified Linear Unit (ReLU) function, defined as

$$f(x)=\max(0,x). \tag{5.2}$$

The use of ReLU and its variants (such as SiLU in the latest versions of YOLO) effectively addresses the problem of gradient vanishing and accelerates model convergence during training.

In modern object detector architecture design, two conceptual strategies are distinguished [14]:

1. Two-stage detectors: R-CNN family. The detection process is divided into the generation of regional proposals and their subsequent classification. Despite their high accuracy, such systems are computationally intensive, which limits their application in real-time tasks [15].

2. One-stage detectors: YOLO, SSD, RetinaNet. These models treat detection as a single regression problem, predicting the coordinates of the bounding boxes and class probabilities in a single pass through the network. This approach is the dominant one in aerial reconnaissance systems due to its high throughput and the ability to deploy it on UAVs with limited computational resources.

However, the detection of small-scale targets remains a "bottleneck" almost exclusively for deep neural networks. The main reason lies in the loss of spatial information that occurs during the sequential downsampling of feature maps. To mitigate this phenomenon, modern architectures employ multi-scale prediction mechanisms and hierarchical feature combination structures (Feature Pyramid Networks, FPN). This allows the network to aggregate contextual information from deep layers with the high spatial resolution of shallow layers.

The evolution of the YOLO family of models has led to the emergence of architectures that demonstrate high robustness to challenging observation conditions: low contrast, spectral blending of the object with the background, and geometric distortions. Within the scope of this study, special attention is paid to a comparative analysis of the YOLOv8 and YOLOv11 models. The choice of these iterations is due to their technological sophistication: the integration of anchor-free detection methods, improved feature extraction blocks, and optimized loss functions, which together open up new prospects for the automated interpretation of aerial reconnaissance data under conditions of active countermeasures and camouflage [16].

5.4 YOLO-based architectures for aerial object detection

Models from the YOLO family are among the most widely used object detection algorithms in modern computer vision systems. Their popularity stems from their combination of high image processing speed and reasonably high object recognition accuracy. Unlike two-stage detectors, such as Faster R-CNN, which first generate region proposals and then classify them, YOLO models

perform object localization and classification in a single pass through the neural network.

The main idea behind the YOLO algorithm is to divide the input image into a regular grid. Each cell of this grid is responsible for predicting objects, which centers are within the corresponding area. For each cell, the neural network predicts the coordinates of the bounding box, the probability of an object's presence, and the probability that the object belongs to a specific class.

The number of parameters predicted by the model can be described as follows

$$S \times S \times (B \cdot 5 + C), \tag{5.3}$$

where S – the size of the grid into which the image is divided; B – the number of predicted frames for each cell; C – the number of object classes.

Each bounding box is described by five parameters: center coordinates x and y, width w, height h, and a confidence value, which characterizes the probability of an object being present within the bounding box.

The architecture of modern YOLO models consists of three main components: backbone, neck, and head. The backbone is used to extract features from the input image using convolutional layers. In this block, feature maps of various levels of abstraction are formed. The neck is designed to combine features at different scales, allowing for more effective detection of both large and small objects. Structures such as Feature Pyramid Networks (FPN) or Path Aggregation Networks (PAN) are often used for this purpose. The head performs the actual prediction of bounding box coordinates and object classes.

To evaluate the quality of object localization, the Intersection over Union (IoU) metric is used, which determines the degree of overlap between the predicted bounding box and the actual bounding box of the object

$$IoU = \frac{A_{rea}\left(B_{pred} \cup B_{gt}\right)}{A_{rea}\left(B_{pred} \cap B_{gt}\right)}, \tag{5.4}$$

where B_{pred} – specified frame; B_{gt} – ground truth frame.

The basic building block of the architecture is the convolutional layer (conv), which structure is shown in **Fig. 5.1**. It consists of a 2D convolutional layer (Conv2d), a batch normalization layer (BatchNorm2d), and a SiLU activation function. This combination allows for the effective extraction of spatial features from the image and stabilizes the neural network training process.

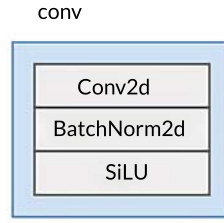


Fig. 5.1 Structure of the convolution block (Conv)
Source: created by the authors

To improve the efficiency of feature processing in the network, bottleneck modules (b_elem) are used. They consist of a sequence of convolutional layers and ensure efficient information transfer between the layers of the neural network. The structure of this block is shown in **Fig. 5.2**.

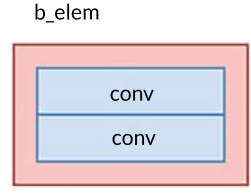


Fig. 5.2 Structure of the bottleneck block (b_elem)
Source: created by the authors

One of the key modules in modern YOLO models is the C2f block, which combines several bottleneck elements with additional convolutional layers. This structure improves feature transfer between layers and enhances the training efficiency of deep neural networks. The architecture of the C2f module is shown in **Fig. 5.3**.

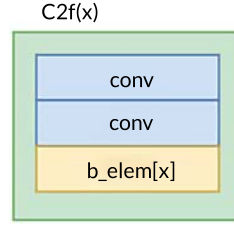


Fig. 5.3 Architecture of the C2f module
Source: created by the authors

The final component of the architecture is the detection head, which predicts the coordinates of bounding boxes and object classes.

This block uses information from the previous network layers to generate the final detection results.

The structure of this component is shown in **Fig. 5.4**.

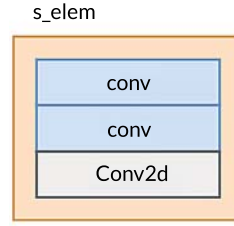


Fig. 5.4 Structure of the detection head element
Source: created by the authors

The combination of these structural elements forms the complete architecture of the YOLO model, which includes feature extraction blocks and an object detection block. The overall structure of the model's architecture is shown in **Fig. 5.5**.

In this study, two state-of-the-art architectures – YOLOv8 and YOLOv11 – were used to detect small and occluded objects in aerial reconnaissance images. The YOLOv8 model is one of the most widely used architectures for object detection tasks and demonstrates a good balance between accuracy and processing speed. However, further development of the YOLO family of models led to the emergence of the newer YOLOv11 architecture, which is characterized by improved computational efficiency and increased detection accuracy [8].

In addition to improvements in the neural network architecture, detection performance depends significantly on image preprocessing methods and the structure of the training dataset.

This study also investigated the impact of various data preparation approaches, including the use of padding to preserve image proportions and the optimization of the number of classes in the dataset.

The analysis conducted allows to evaluate the effectiveness of modern YOLO architectures in automated aerial reconnaissance image analysis tasks and identify the most promising approaches for detecting small-sized and camouflaged objects.

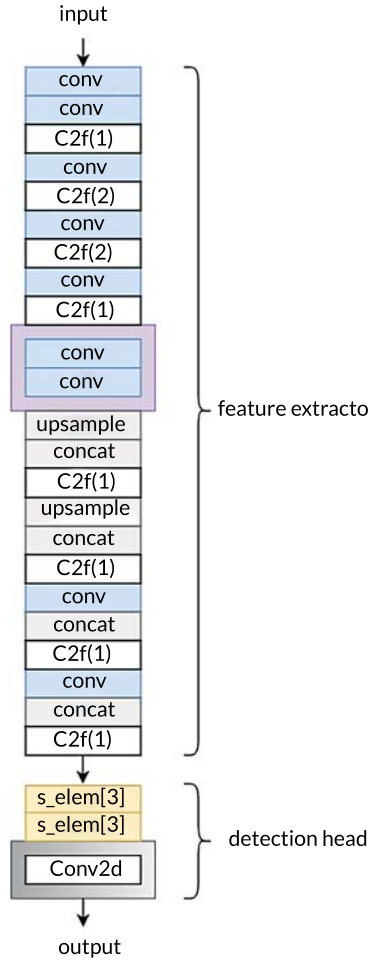


Fig. 5.5 Overall architecture of the YOLO model
Source: created by the authors

5.5 Experiment using YOLOv8 for small object detection

To investigate the effectiveness of modern methods for detecting small and occluded objects, an experiment was conducted using the YOLOv8 model. This model

belongs to the modern generation of single-stage object detectors and is widely used in computer vision tasks due to its combination of high image processing speed and sufficiently high detection accuracy [13].

In the first stage, a training dataset was created based on aerial reconnaissance video footage. The video recordings were divided into individual frames, which were used as images for further training of the neural network. For each image, object annotation was performed, during which the coordinates of the bounding boxes and the corresponding object classes were determined.

Analysis of the dataset structure is a crucial step in experiment preparation, as class distribution can affect the quality of model training. The distribution of objects by class is shown in **Fig. 5.6**.

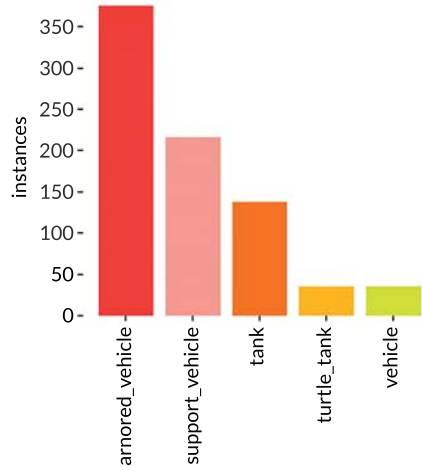


Fig. 5.6 Histogram of class distribution in the training dataset
Source: created by the authors

Further analysis of the positions of objects in the images allows to assess the spatial distribution of objects in the dataset. **Fig. 5.7** shows the distribution of the positions of the frame centers and their sizes.

After preparing the dataset, the YOLOv8 model was trained. During training, the neural network gradually optimizes its parameters by minimizing the loss function and improving object detection accuracy.

The trends in the model's key quality metrics over the training epochs are shown in **Fig. 5.8**.

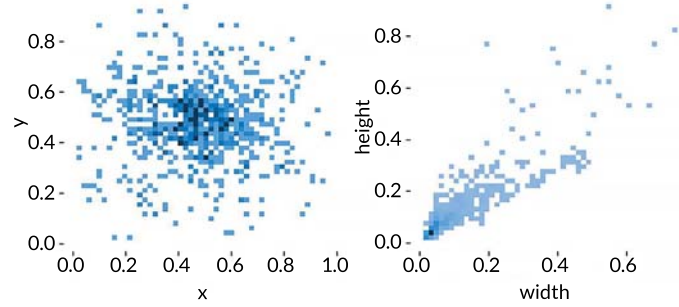


Fig. 5.7 Distribution of positions (left) and dimensions (right) of the limiting frames
Source: created by the authors

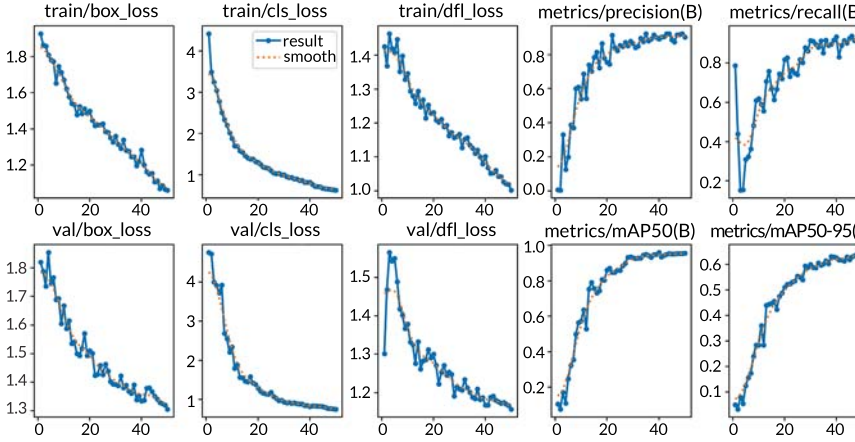


Fig. 5.8 Changes in the model's quality metrics over training epochs
Source: created by the authors

Standard object detection metrics are used to evaluate the model's performance. One of the key metrics is Precision, which measures the proportion of correctly detected objects among all detections

$$Precision = \frac{TP}{TP + FP}, \tag{5.5}$$

where *TP* (True Positive) refers to correctly detected objects, and *FP* (False Positive) refers to false detections.

Another important metric is Recall, which shows what proportion of actual objects the model was able to detect

$$Recall = \frac{TP}{TP + FN}, \tag{5.6}$$

where *FN* (False Negative) refers to objects that the model failed to detect.

To comprehensively evaluate detection quality, the *F*1-score is used, which combines the Precision and Recall metrics

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}. \tag{5.7}$$

Using this metric provides a comprehensive assessment of the model's performance, taking into account both detection accuracy and recall.

For a more detailed analysis of the model's performance, curves were plotted showing the dependence of key detection metrics on the model's confidence level (**Fig. 5.9**).

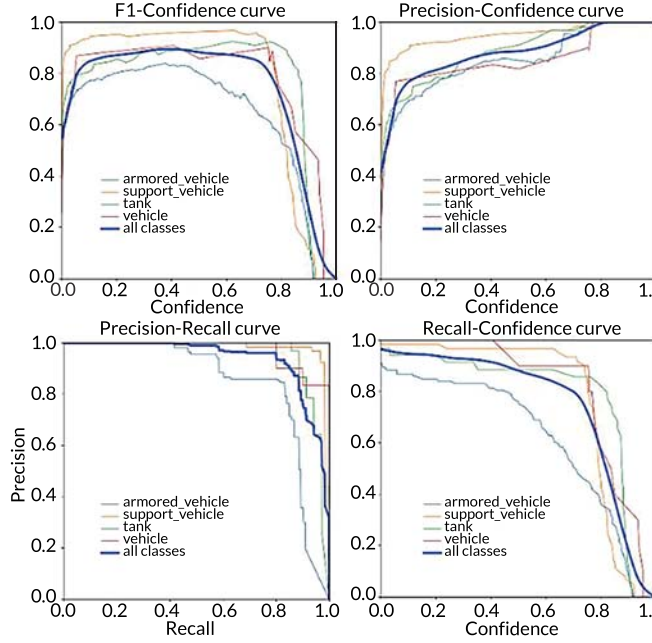


Fig. 5.9 Dependency curves for key detection metrics in the YOLOv8 model
Source: created by the authors

These graphs allow to assess the impact of the confidence threshold on detection accuracy and completeness, as well as to determine the optimal threshold value for the practical application of the model.

As can be seen from the graphs, as the confidence threshold increases, detection accuracy (Precision) improves, but at the same time, the completeness of object detection (Recall) decreases. The Precision-Recall curve illustrates the relationship between these two metrics. The maximum *F*1-score corresponds to the optimal balance between precision and recall. A confusion matrix is used to evaluate the classification quality of different object types. It allows to determine which object classes are most frequently misclassified by the model. The normalized confusion matrix is shown in **Fig. 5.10**.

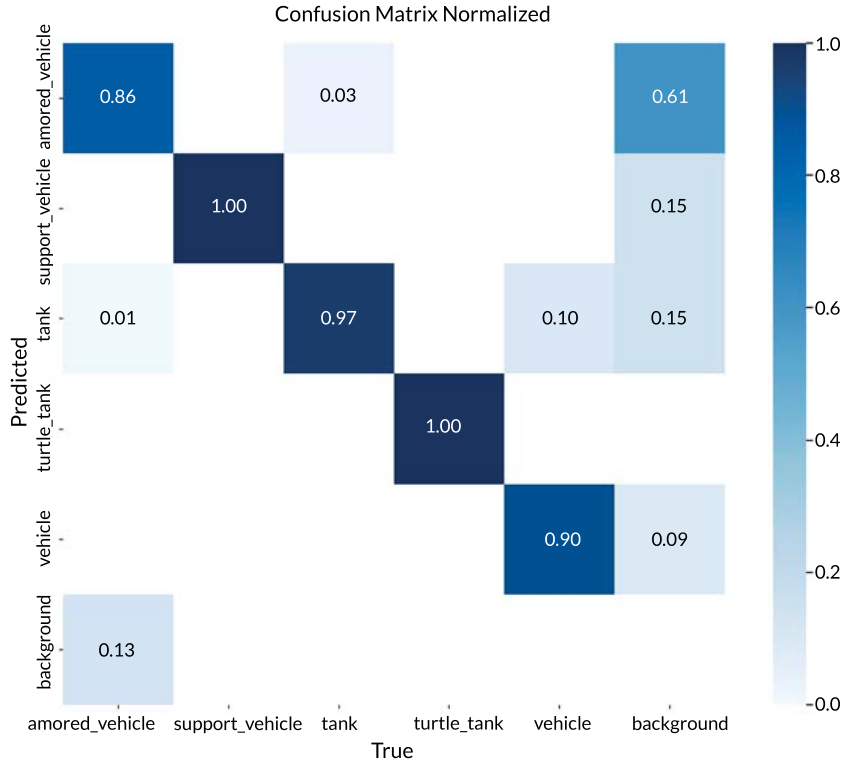


Fig. 5.10 Normalized confusion matrix for the YOLOv8 model
Source: created by the authors

The results showed that the YOLOv8 model is capable of effectively detecting objects in aerial reconnaissance images. At the same time, a number of limitations were identified regarding the detection of small and partially occluded objects, which is attributed to their low pixel count in the image and complex background conditions.

The experimental results were used as a baseline for further comparison with the newer YOLOv11 architecture, which incorporated additional optimizations in image preprocessing and dataset structure.

5.5.1 Experiment using YOLOv11 for small object detection

Following the baseline experiment using the YOLOv8 model, a follow-up study was conducted using the newer YOLOv11 architecture. The primary objective of this experiment was to improve the detection accuracy of small and occluded objects, as well as to optimize the model training process.

The new experiment used the same dataset as the previous study, but several changes were made to the data preparation and model configuration. Specifically, instead of resizing images to a fixed size, the padding method was used, which involves adding blank areas to the image to achieve the required size without altering the aspect ratio. When using the padding method, the image is scaled to a specified size without changing the aspect ratio, after which empty areas are added to the image. Let the original image have a size of $W \times H$, and the required size of the neural network's input image be $S \times S$.

The scaling factor is defined as

$$r = \min\left(\frac{S}{W}, \frac{S}{H}\right). \tag{5.8}$$

After resizing, the new image dimensions are defined as

$$W' = rW, H' = rH. \tag{5.9}$$

The size of the padding area is calculated as follows:

$$p_x = \frac{S - W'}{2}, p_y = \frac{S - H'}{2}, \tag{5.10}$$

where p_x – horizontal padding; p_y – vertical padding.

In addition, the class structure in the dataset was optimized. Reducing the number of classes improves classification confidence, as the neural network focuses on a smaller number of object categories. This also reduces image processing time and increases the speed of the detection system.

During model training, a significant reduction in training time was observed. While in the previous experiment using YOLOv8, model training took more than two hours, in the case of YOLOv11, the full training cycle was completed in approximately 30 minutes.

This indicates the increased efficiency of the new architecture and the optimization of the neural network training process.

The dynamics of changes in the model's key quality metrics over training epochs are shown in **Fig. 5.11**.

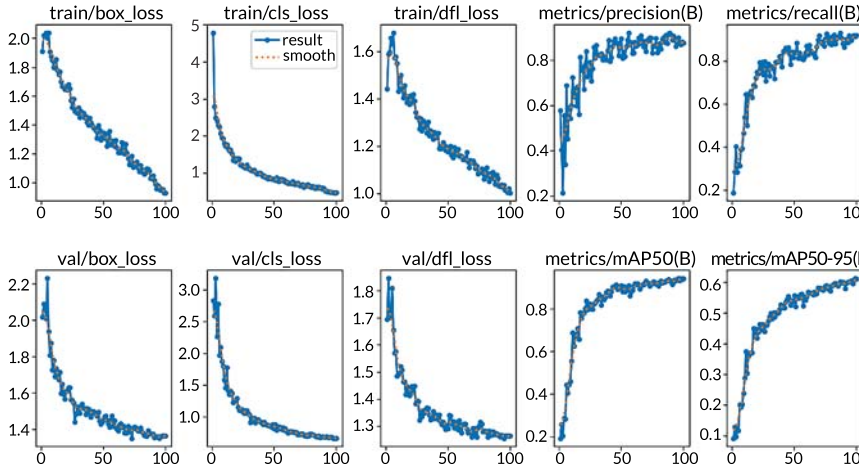


Fig. 5.11 Changes in the key performance metrics of the YOLOv11 model during training
Source: created by the authors

To analyze the model's performance in greater detail, it is possible to plot curves showing how key detection metrics vary with the confidence threshold. These metrics include Precision, Recall, and *F*1-score, which are widely used to evaluate the quality of object detection systems.

The Precision-Confidence, Recall-Confidence, Precision-Recall, and *F*1-Confidence graphs allow to evaluate the model's behavior as the confidence threshold

changes and to determine the optimal balance between detection accuracy and completeness (**Fig. 5.12**).

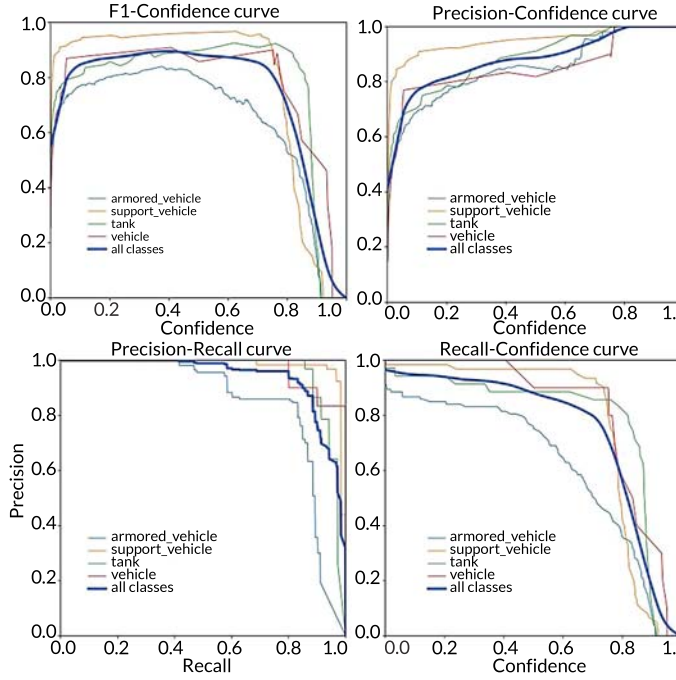


Fig. 5.12 Performance curves for the YOLOv11 model
Source: created by the authors

An analysis of the graphs shows that the model exhibits stable metric values across a wide range of confidence thresholds. As the threshold increases, the Precision metric rises, indicating a decrease in the number of false detections. At the same time, the Recall metric gradually decreases, as some objects may be missed when the threshold is set too high.

For a more detailed analysis of the classification results, a confusion matrix was constructed, which allows for the evaluation of the accuracy of identifying individual object classes. The normalized confusion matrix is shown in **Fig. 5.13**.

As shown in the confusion matrix, most objects are classified correctly, indicating the model's high accuracy. The small number of classification errors is due to the similarity of certain object types and challenging shooting conditions.

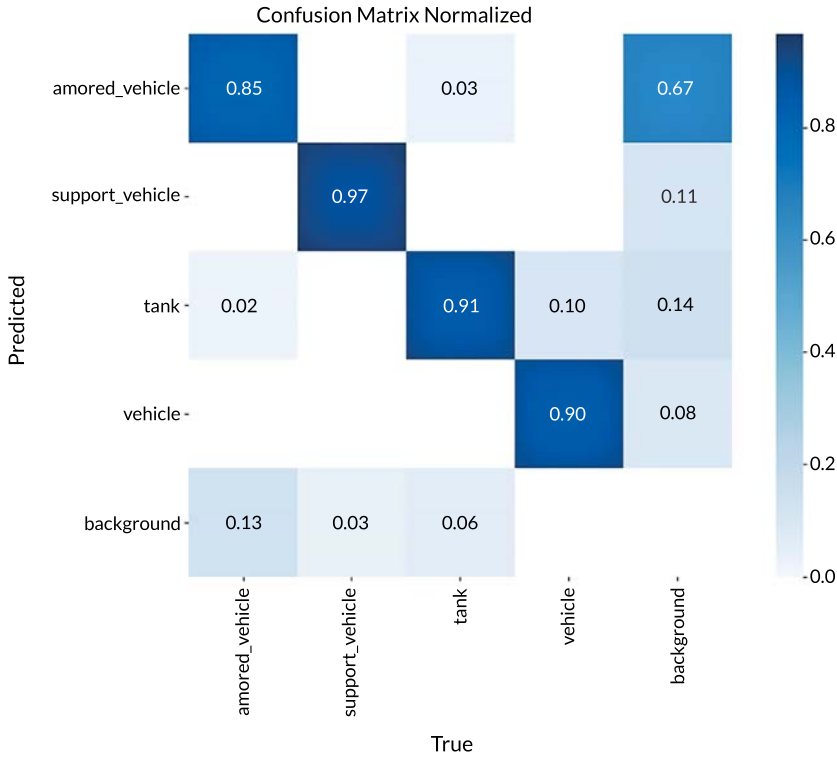


Fig. 5.13 Normalized confusion matrix for the YOLOv11 model
Source: created by the authors

The results demonstrate that using the YOLOv11 model in combination with optimized data preprocessing improves the detection performance of small objects and reduces model training time. This makes this approach promising for use in automated analysis systems for aerial reconnaissance images.

5.6 Comparative analysis of YOLOv8 and YOLOv11

To evaluate the effectiveness of the proposed approach, a comparative analysis was conducted between the YOLOv8n and YOLOv11s models, which were used to detect small and partially hidden objects in photos and videos.

In the second experiment, several changes were made to the data preprocessing and model tuning. First, the image preprocessing method was modified. While in the first experiment, images were resized to the required dimensions, in the second experiment, padding was used, i.e., adding empty areas to the image to achieve the required size without changing its proportions. This approach preserves the geometric characteristics of objects and avoids their deformation, which positively impacts detection quality.

Second, the study utilized the newer YOLOv11s neural network architecture. Version 11's model utilizes computational resources more efficiently, as evidenced by network training speed and data processing speed.

Furthermore, the dataset structure was optimized by reducing the number of classes. This significantly increased the probabilities of class detection, particularly for the "tank" class. The YOLOv8n model took over two hours to train, while the YOLOv11s model was trained in approximately 30 minutes, indicating significant optimization of the training process.

A comparison of the key characteristics of the models is presented in **Table 5.1**.

Table 5.1 Comparison of model specifications

Parameter	YOLOv8n	YOLOv11s
Network architecture	YOLOv8n	YOLOv11s
Image preprocessing	Image resizing	Padding with preserved aspect ratio
Dataset size	> 750 images	> 750 images
Number of classes	initial dataset	optimized dataset
Training time	> 2 hours	< 30 minutes
Precision	≈ 0.90	≈ 0.91–0.93
Recall	≈ 0.88–0.90	≈ 0.90–0.92
mAP@0.5	≈ 0.95	0.939
Detection confidence	moderate	higher

Source: created by the authors

For a comparative analysis, **Fig. 5.14** shows the results of object detection obtained using the YOLOv8n and YOLOv11s models. The examples provided illustrate how the models perform when detecting small objects in aerial reconnaissance images.

Fig. 5.14 Comparison of detection results: 1 – YOLOv8n; 2 – YOLOv11s
Source: created by the authors

5.7 Conclusions

The problem of detecting small and camouflaged objects is highly relevant. This chapter demonstrates the potential of using YOLO family models for this task, as they offer a combination of high speed and sufficient accuracy.

This study analyzed the architectural features of YOLO models and demonstrated their structural elements. An experimental performance evaluation of the YOLOv8n and YOLOv11s models was also conducted. For training and testing, a dataset was created based on 25 videos totaling over 1 hour, from which over 750 images were extracted and annotated.

A baseline experiment using YOLOv8n confirmed the feasibility of effective object detection in photo and video data. However, the results demonstrated the need for further refinement of the approach, particularly with regard to image preprocessing, class set structure, and model architecture.

In the second experiment, the YOLOv11s model was applied, and the data preparation procedure was improved. Instead of directly resizing images, padding was used, preserving the proportions of objects and reducing geometric distortions. Furthermore, the class structure was optimized, which positively impacted classification confidence and detection speed.

The results showed that the YOLOv11s model provides high detection performance: AP = 0.848 for the armored_vehicle class, 0.983 for support_vehicle, 0.956 for tank, and 0.968 for vehicle, while the overall AmAP@0.5 is 0.939. Model training time was reduced from over two hours for YOLOv8n to approximately 30 minutes for YOLOv11s, indicating a significant increase in computational efficiency.

Thus, the results of this study confirm the promise of using modern YOLO family architectures for automated data analysis. The combination of improved data preprocessing methods, class structure optimization, and the use of new detection models allows for increased reliability in detecting small and camouflaged objects in challenging surveillance environments. The main limitations and assumptions of the proposed approach are summarized in **Table 5.2**.

These limitations and assumptions define the directions for future research and indicate the need for validation on larger and more diverse datasets, as well as testing under real-world operating conditions.

Prospects for further research include expanding the training dataset and further adapting the models to practical conditions in monitoring and reconnaissance systems. It is also proposed to increase data augmentation at subsequent stages of research.

Table 5.2 Main limitations and assumptions of the study

Category	Description
Limitation	The dataset is relatively limited in size and diversity, as it is based on publicly available video data
Limitation	The model was not evaluated under extreme conditions such as severe illumination changes, weather effects, or sensor noise
Limitation	Motion blur, rapid viewpoint changes, and real-time communication delays were not explicitly considered
Limitation	Limited validation under real UAV operational conditions
Assumption	Image acquisition conditions are assumed to be relatively stable
Assumption	Preserving object geometry using padding is considered more important than potential loss of contextual information
Assumption	Reducing the number of classes improves detection performance
Assumption	The selected YOLOv11 architecture is appropriate for the defined task

Source: created by the authors

Conflict of interest

The authors declare that there is no conflict of interest in relation to this paper, as well as the published research results, including the financial aspects of conducting the research, obtaining and using its results, as well as any non-financial personal relationships.

Financing

The research has been funded by National Research Foundation of Ukraine (accessed on 24 June 2025) within Project No. 2025.06/0037 "A system for detecting and recognizing camouflaged and small objects based on the use of modern computer vision technologies" (2025–2026).

Data availability

Data will be made available on reasonable request.

Use of artificial intelligence statements

The authors confirm that they did not use artificial intelligence technologies when creating the current work.

Authors' contributions

Dmytro Krytskyi: Approved the research concept, Supervised the study, Contributed to the research design, Participated in the interpretation of the results, and Critically revised the manuscript.

Elvira Kaidan: Prepared and annotated the dataset, Assisted in conducting the experiments, Contributed to the software development, and Preparation and formatting responsibilities of the manuscript materials.

Artem Chekhovsky: Contributed to the software implementation, Validation of the obtained results, and Manuscript editing.

References

1. Pan, Yu., Li, L., Qin, J., Chen, J.-J., Gardoni, P. (2024). Unmanned aerial vehicle-human collaboration route planning for intelligent infrastructure inspection. Computer-Aided Civil and Infrastructure Engineering, 39 (14), 2074–2104. https://doi.org/10.1111/mice.13176
2. Zeng, B., Gao, S., Xu, Y., Zhang, Z., Li, F., Wang, C. (2024). Detection of Military Targets on Ground and Sea by UAVs with Low-Altitude Oblique Perspective. Remote Sensing, 16 (7), 1288. https://doi.org/10.3390/rs16071288
3. Hwang, K.-S., Ma, J. (2024). Military camouflaged object detection with deep learning using dataset development and combination. The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology, 23 (1), 67–78. https://doi.org/10.1177/15485129241233299
4. Akyon, F. C., Onur Altinuc, S., Temizel, A. (2022). Slicing Aided Hyper Inference and Fine-Tuning for Small Object Detection. 2022 IEEE International Conference on Image Processing (ICIP), 966–970. https://doi.org/10.1109/icip46576.2022.9897990
5. Li, Y., Li, Q., Pan, J., Zhou, Y., Zhu, H., Wei, H., Liu, C. (2024). SOD-YOLO: Small-Object-Detection Algorithm Based on Improved YOLOv8 for UAV Images. Remote Sensing, 16 (16), 3057. https://doi.org/10.3390/rs16163057

6. Han, Y., Guo, J., Yang, H., Guan, R., Zhang, T. (2024). SSMA-YOLO: A Lightweight YOLO Model with Enhanced Feature Extraction and Fusion Capabilities for Drone-Aerial Ship Image Detection. Drones, 8 (4), 145. https://doi.org/10.3390/drones8040145
7. Yang, C., Shen, Y., Wang, L. (2025). EMFE-YOLO: A Lightweight Small Object Detection Model for UAVs. Sensors, 25 (16), 5200. https://doi.org/10.3390/s25165200
8. Kaidan, E. (2026). Improved object recognition methods in UAVS based on YOLO. Propylaea of Law and Security, 8, 218–221. https://doi.org/10.32620/pls.2025.8.55
9. Ali, M. L., Zhang, Z. (2024). The YOLO Framework: A Comprehensive Review of Evolution, Applications, and Benchmarks in Object Detection. Computers, 13 (12), 336. https://doi.org/10.3390/computers13120336
10. Krytskyi, D., Tkachov, I., Pyvovar, M., Karatanov, S., Krikun, A. (2025). Detecting Objects in Photos and Videos Obtained by Aerial Reconnaissance from UAVs. Integrated Computer Technologies in Mechanical Engineering – 2024. Cham: Springer, 270–283. https://doi.org/10.1007/978-3-031-94845-9_23
11. Kritsky, D., Kaidan, E., Tkachov, I., Lukin, V. (2025). Automatic detection of hidden objects using uavs: modern neural network approaches. Herald of Khmelnytskyi National University. Technical Sciences, 359 (6.2), 193–204. https://doi.org/10.31891/2307-5732-2025-359-98
12. Mi, Q., Chao, J., Chen, A., Zhang, K., Lai, J. (2026). YOLO11s-UAV: An Advanced Algorithm for Small Object Detection in UAV Aerial Imagery. Journal of Imaging, 12 (2), 69. https://doi.org/10.3390/jimaging12020069
13. Makarichev, V., Tsekhmystro, R., Lukin, V., Krytskyi, D. (2025). Performance Improvement of Vehicle and Human Localization and Classification by YOLO Family Networks in Noisy UAV Images. Information, 16 (12), 1087. https://doi.org/10.3390/info16121087
14. Chaini, C., Jha, V. K., Rajnish, K. (2026). A comparison of single-stage and two-stage based crater detectors on the lunar surface. Earth Science Informatics, 19 (2). https://doi.org/10.1007/s12145-025-02066-7
15. Gu, H., Wu, J., Huang, H. (2026). CASA-RCNN: A Context-Enhanced and Scale-Adaptive Two-Stage Detector for Dense UAV Aerial Scenes. Drones, 10 (2), 133. https://doi.org/10.3390/drones10020133
16. Stets, S. (2025). Analysis of accuracy and speed of vehicle detection using neural networks YOLOv8 and YOLOv11. Herald of Khmelnytskyi National University. Technical Sciences, 357 (5.2), 123–130. https://doi.org/10.31891/2307-5732-2025-357-74

CHAPTER 6

Visual pattern recognition in navigation simulator interfaces: a method for automatic reconstruction of vessel motion parameters

Yevgeniy Kalinichenko
Oleksandr Koliesnik
Oleg Safyan

Abstract

In this chapter, it is presented a method for automatically determining the navigation and control parameters of a vessel, which is based on reading the simulator screen without access to internal telemetry or system data. The presented approach considers the display of the simulator not only as an interaction tool for the operator, but also as a visual representation of the state of the internal model, because the navigation parameters are coded using graphic indicators that create a graphical user interface, which methodologically can be understood as a structured visual environment where stable areas correspond to certain parameters.

By automatically identifying these areas and interpreting their contents, the method determines synchronized time series that describe the vessel's movement. Once set up, the system can read numerical and graphical indicators and convert them into data sequences up to 10 Hz, allowing quantitative information on vessel movements to be obtained directly from the user interface image without software integration or access to internal modules.

The experimental validation was performed on a fully functional bridge simulator which user interface displays all navigation and engineering parameters, and its results showed that the most important navigation parameters can be reconstructed very reliably using only the interface image, while the reconstructed time series supports the correct sequence of events. The average absolute error is a maximum of 1° for the course and remains below 2% for propeller revolutions, which is sufficient for maneuver analysis and trajectory studies.

A key advantage of this approach is its independence from the internal architecture of specific simulation systems. This allows the method to be applied on various platforms, even where access to internal data is restricted or unavailable.

 DOI: 10.21303/978-9908-845-05-0.ch6

The method is currently protected by two patent applications filed with the National Intellectual Property and Innovation Agency of Ukraine.

Overall, the results demonstrate that the visual interface of a navigation simulator can serve as an independent and universal data source for investigating ship dynamics and supporting experimental studies.

Keywords

Image recognition, navigation simulator, interface analysis, ship motion parameters, automated data collection, ship dynamics, bridge simulator.

6.1 Introduction

Over the past decades, digital simulation technologies have significantly expanded the role of navigational simulators in maritime research and training. Modern full-mission bridge simulators reproduce not only the visual layout of a ship's wheelhouse, but also the operational logic of navigation systems and ship control. As a result, they provide a controlled environment in which ship dynamics, environmental influences, and operator decisions can be studied in conditions that are as close as possible to real navigational practice. At the same time, a practical difficulty appears when such studies require access to reliable navigation data, especially when internal simulator outputs are limited or unavailable.

In this chapter, this problem is addressed by proposing a method for extracting navigation and control parameters directly from the visual interface of the simulator. The goal is to develop an approach that makes it possible to reconstruct vessel motion without relying on internal telemetry, using pattern recognition methods applied to elements of the graphical interface. Despite these advantages, obtaining experimental data from simulators remains a technically challenging task.

In many cases, researchers have to rely on internal telemetry channels, proprietary service logs, or specialized application programming interfaces to access parameters describing the ship's motion. Such solutions often depend on the specific architecture of the simulator software. As a result, experimental procedures developed for one system cannot always be reproduced on another platform. When the internal structure of the simulator is closed or poorly documented, the parameters necessary for analysis may remain unavailable to the researcher.

The general theme of this monograph, pattern recognition in surveillance and diagnostic systems, suggests a different approach to solving this problem. What matters here is not the recognition itself. It is how the simulator is treated as a data source. The graphical interface of a navigation simulator continuously displays

a large amount of operational information about the state of the vessel. From a methodological point of view, it can be considered not only as a tool for interaction with a person, but also as a visual representation of the internal state of the model.

In other words, the simulator screen can be considered as a structured visual environment in which navigation parameters are encoded using graphic indicators.

Analysis of these indicators using pattern recognition methods allows obtaining quantitative information about the movement of the vessel without direct access to the internal software modules of the simulator.

The method described in this chapter is based on this idea.

Instead of using internal telemetry streams, the proposed approach analyzes the graphical interface of the simulator itself. The display is considered as an external projection of the internal state of the simulator, and the graphic elements that appear on the screen are interpreted as a representation of navigation and control parameters.

By automatically identifying these elements, the method reconstructs synchronized time series describing the movement of the simulated vessel.

The experimental verification of the approach was carried out using a full-mission bridge simulator.

The technical solution presented here is currently protected by two pending patent applications filed with the Ukrainian National Office for Intellectual Property and Innovation (UANIPIO): application for invention No. a202600852 and application for utility model No. u202600851.

The rest of the chapter is organized as follows. Section 6.2 reviews previous research on marine simulators and visual pattern recognition methods. Section 6.3 examines the navigation simulator interface as a structured visual environment and outlines the conceptual framework of the proposed approach. Section 6.4 describes the architecture of the method and its principles of operation. Section 6.5 presents the results of the experimental verification. The final section summarizes the main results of the study.

6.2 Related works

Over the past twenty years, research on navigation simulators has been significantly intensified. This is due to the fact that simulators have begun to play an important role not only in training but also in scientific research into ship control.

Z. Munim et al. show in their systematic review that simulator-based research is becoming more methodologically structured, especially in terms of scenario

evelopment, data measurement and analysis [1]. But their work focuses mainly on how such research is conducted, rather than how navigational data can be obtained independently of the specific architecture of the simulator.

C. Sellberg considers simulators primarily as a teaching and evaluation tool and clearly demonstrates their value for reproducing complex navigational situations [2]. However, the simulator is considered primarily as a pedagogical environment, while the movement of the vessel itself is not distinguished as an independent object of quantitative analysis.

H. Tasher et al. investigate marine simulators from technical, training and organizational perspectives and point to the increasing complexity of their systems [3]. This paper is useful for understanding a broader evaluation system, but it also proceeds from the assumption of internal simulator data availability, and does not address alternative ways of obtaining navigation parameters without direct access to the system.

Modern simulators reproduce the operation of the ship's bridge. They combine the movement of the vessel, the influence of the environment and the actions of the boatmaster. Thanks to this, simulators gradually became not just educational tools, but full-fledged experimental sites where navigation processes can be studied in controlled conditions.

Usually, such research is focused on three directions: creation and evaluation of the simulators themselves, analysis of the behavior of ship drivers and modeling of ship movement.

There are many works devoted to the use of simulators in education and science. Reviews show that simulators are widely used to study decision-making, maneuvering, and situation assessment in difficult environments.

These reviews also show a common limitation, namely that the issue of obtaining data usually remains within the simulator environment itself and is rarely considered as a separate methodological problem.

The advantage of simulators is that they allow to safely reproduce rare or dangerous situations. This makes it possible to obtain data that is difficult or impossible to collect in a real voyage.

At the same time, such studies have their own difficulties. It is necessary to correctly create scenarios, determine what exactly to measure, and correctly interpret the received data.

Another very important difficulty in comparative or repeated studies is the dependence on the internal outputs of the simulator, its service logs or its own interfaces, which prevents even well-planned experiments from being transferred from one platform to another.

In most cases, the simulator is seen as an environment where the behavior of the vessel can be observed in different conditions. In parallel, the behavior of the boatmaster is also studied: how it perceives information, assesses the situation and makes decisions during its change. In recent years, special attention has been paid to the so-called non-technical skills, the level of cognitive load and the peculiarities of the distribution of visual attention.

C. Hetherington et al. emphasize the importance of the human element in maritime safety and show that decision-making and cognitive factors often determine the outcome of navigational situations [4]. At the same time, their work does not consider the problem of obtaining quantitative data on vessel motion.

F. Saeed et al. propose a method for assessing non-technical skills in a bridge simulator using evidential reasoning [5]. Their approach allows evaluating operator performance, but the vessel itself is treated mainly as a background for such assessment.

O. Atik and O. Arslan use eye-tracking methods to analyze how operators interact with navigation information [6]. This makes it possible to study visual attention in detail, although the parameters of vessel motion are not analyzed independently.

V. Ronca et al. apply neurophysiological monitoring to assess operator state during emergency training in a full-mission simulator [7]. Their results give additional insight into cognitive load, but again focus on the operator rather than on vessel dynamics.

To study these aspects, various methods of observing the operator's activities are used. The most common are eye tracking systems, behavioral coding methods, as well as neurophysiological monitoring, which allows assessing the operator's state while performing navigation tasks in a simulation environment. Such approaches provide valuable information about the role of the human element in the process of controlling a vessel. However, in most such studies, the dynamics of the vessel itself are secondary. In most cases, vessel motion is considered mainly as a consequence of operator decisions, while the possibility of analyzing movement parameters as an independent object of quantitative study is addressed much less often.

Another notable line of research is related to the mathematical and physical description of ship maneuvering. The main attention in such works is paid to the development and refinement of hydrodynamic models that allow describing the motion of a ship under the influence of external factors – wind, waves and currents. As a rule, it is about constructing systems of differential equations that reflect the interaction between the engine, steering control and external forces acting on the hull.

After formalizing the model, its parameters are refined taking into account experimental measurements or data obtained during modeling. Such models are widely used to analyze the maneuvering characteristics of a ship, predict its trajectory and develop algorithms for automatic motion control.

For example, V. Frett et al. analyze maneuvering data from simulators and show that it can be used to study ship dynamics, but they rely completely on data recorded inside the simulator [8].

H. Li et al. look at uncertainty in maneuver simulations and highlight the need for accurate parameters, although their approach still assumes access to model or simulator data [9].

P. Pires da Silva studies how sensitive maneuvering models are to changes in parameters, showing how much results can vary. This helps to understand the models better, but does not explain how these parameters can be obtained in practice [10].

H. He et al. describe a simulator used for testing automatic control algorithms, where the system is treated as closed and data can only be accessed through internal tools [11].

H. Yasukawa and Y. Yoshimura present the MMG method, which is widely used to predict ship maneuvering. It provides a clear modeling framework, but assumes that all required parameters are already known or measured [12].

J. Lee et al. propose an improved model that includes wave effects using a two-time-scale approach. This gives a more realistic description, but still depends on predefined input data [13].

D. Kim et al. develop a CFD-based model for maneuvering in currents. It provides detailed results, but requires high computational effort and carefully prepared input data [14].

At the same time, the source of data for such studies is most often either the mathematical models themselves or the internal telemetry of navigation simulators. As a result, the task of reconstructing navigation parameters from external observations without referring to the internal data of the system, as a rule, remains beyond the scope of these works.

In practice, navigation data is usually taken straight from the simulator software. This is done through logs, telemetry, or built-in export tools that record parameters like course, speed, and control actions. The data is accurate, but it creates a dependency.

Once a study uses internal simulator data, it becomes tied to that specific system. In practice, this is not always convenient. Access can be limited, poorly documented, or different from one simulator to another. Because of this, even a good experiment can be hard to repeat on another platform. Often, extra integration or custom tools are needed.

In other fields, things are handled differently. Computer vision is already used to read data from screens in industrial monitoring, diagnostics, and control systems.

For example, R. Smith describes the Tesseract OCR engine, which can read text directly from images and is widely used to extract data from visual interfaces [15]. This shows that even simple visual elements can be turned into structured data without internal access.

R. Szeliski gives a general overview of computer vision methods, including how to detect objects, recognize them, and track them in images [16]. These ideas can be applied to reading interface elements on a screen.

T. Wuest et al. show that in many industrial systems, data is taken from external observations rather than internal sources, and machine learning is used to process visual and sensor data [17].

M. Rehman et al. review computer vision methods in construction, where system states are reconstructed from images and video, showing that visual data alone can be enough to track complex processes over time [18].

The screen is treated as a source of information: indicators are located, their values are read, and the system state is reconstructed without using internal data.

This approach is common in industry, but in maritime simulators it is still rare. This creates a clear gap.

On one side, maritime research produces a lot of data about ship motion and human performance – but almost always through internal simulator logs. On the other side, computer vision already has tools that can extract the same kind of data from images – but these methods haven't really been adapted to ship simulators.

So, in most studies, one simple idea is still overlooked: what if the simulator interface itself could be used as an independent and universal source of navigation data?

This observation is the starting point of this work. If to change the simulator screen view, it is possible to see what it does as an operator interface. Indeed, the screen displays the current state of the ship. All instruments and indicators on the screen – steering, engine thrust, rotational speed, ambient parameters – reflect the values of the variables calculated in the simulation. The graphical elements of the interface can therefore be considered as a visible visualization of the internal parameters of the system.

In this interpretation, the interface acts as a structured visual space. In it, different areas of the screen correspond to specific navigation indicators, and parameters are displayed through stable graphic forms. If to analyze these interface elements, it is possible to reconstruct the value of the ship's movement parameters and the time change. This approach allows to receive data directly from the interface image without connecting to the internal telemetry modules of the simulator. Thanks to this, it becomes possible to conduct research even in cases where access to internal software is limited. The method proposed in the work arose at the

intersection of two areas of research. The first is related to the use of marine simulators in scientific experiments. The second is with the development of computer vision and pattern recognition methods. By applying image recognition methods to the interface of the bridge simulator, it is automatically possible to read the readings of the devices and form synchronized time series of navigation parameters directly from the screen image. The analysis of existing studies shows that these two directions developed separately from each other for a long time. Works devoted to navigation simulators mainly focus on training of boatmasters, analysis of operators' behavior and improvement of vessel propulsion models. At the same time, methods of obtaining quantitative information from graphic interfaces of various technical systems are actively developing in the field of computer vision. Despite the apparent closeness of these ideas, their joint use in marine modelling has been virtually unaddressed.

Therefore, the interface of the navigation simulator, although it constantly reflects a large number of parameters of the ship's operation, was not perceived as an independent source of data for analysis for a long time. However, if one considers the screen as a visual representation of the state of the model, it becomes clear that it contains all the necessary information about the dynamics of the vessel. As a result, the task of data collection actually moves from the field of software integration to the field of image analysis. The main task is to find the desired indicators on the screen, determine their location and correctly read the shown values. The possibility of such an approach is explained by the features of modern bridge simulators. Their interfaces usually have a stable structure: the devices are located in fixed areas of the screen, and their appearance during operation practically does not change. At the same time, the values of the indicators change over time in accordance with the models of ship movement. It is this stability that makes it possible to systematically analyze the information displayed on the screen.

6.3 Navigation simulator interface as a structured visual field

The most common approach is the use of internal digital registers and telemetry flows, which are analyzed after the simulation is complete. The trajectory, speed, trajectory and interaction with the navigation objects are removed from the respective data files to evaluate the quality of the maneuvers and compare the scenarios. This approach provides very detailed data and is related to the architecture and registration format of a particular simulation platform (**Fig. 6.1**). Therefore, portability between systems is very limited.

Fig. 6.1 Interface of the navigation simulator, selected during the experiment

The second group focuses on analyzing operator behavior in the simulator. Behavioral markers, non-technical skill scores, eye-tracking data, and cognitive load actions are used to investigate decision-making and situational awareness. In this sense, the dynamics of the ship are perceived as a reflection of the operator's behavior and not as an independent object of quantitative research. The trajectory of a ship is rarely systematically analyzed.

The third study group deals with the mathematical and physical modeling of ship movements using hydrodynamic and hydrodynamic computational equations to refine the predictions. The data from these studies are generated within a mathematical model and require a different external method of data collection.

A significant deviation can be observed in all three methodological study groups: either internal direct access to the simulator data is assumed, or the analysis is limited to the operator's behavior. None of the existing approaches consider the visual layer of the simulator as an independent and universal data source that can provide

reconstructed navigation settings without any software integration. In this context, the research results obtained on one simulation platform cannot easily be reproduced on another platform. Comparative studies of different systems require either rarely occurring standard export interfaces or specialized integration tools. The lack of an independent data collection method limits the scope of simulator research and its application in cases where internal access to the software is limited or unavailable.

As mentioned, the simulator navigation indicators are placed in fixed and predictable areas of the screen. Each indicator has a stable graphic format. It can be a digital display, a dial, an indicator in the form of a strip or a map element. The way information is updated on the screen is an important feature. Since data updates occur sequentially and predictably, automatic screen reading of information and systematic data collection becomes possible.

Consequently, the interface can be thought of as a two-dimensional space with coordinates (x, y) in which each visual element occupies a specific region Z_i. Suppose that $I(x, y, t)$ denotes a function describing the visual state of the interface at time t. Z_i zones correspond to $p_i(t)$ parameters, each of which represents a specific aspect of the ship's condition. Therefore, the internal state of the simulator is described by the state vector $S(t)$, which includes variables of different types – kinematic, dynamic and ecological. The display does not display the status vector completely, but only displays a subset of the variables required for navigation and control. The main purpose of the mapping method is

$$F : I(x,y,t) \rightarrow P(t),$$

where $P(t)$ – a subset of the internal state $S(t)$ that can be recovered from the visual representation. This mapping is not active: some internal variables do not have a direct graphical counterpart, and some interface elements combine several parameters. Therefore, reconstruction is limited to indicators with stable, unambiguous visual representations.

For a method to work stably, it doesn't have to depend exactly on how the interface is displayed. This applies to the screen resolution, scale or graphic settings. Coordinate normalization is used for this. In other words, the surface elements are not described by exact pixels, but by the relative position on the screen. For example, if the area is in the top left corner of the screen, it will stay there even after the resolution has changed (e.g. from 1280 × 1024 to 1920 × 1080).

This property of invariance is important to use the method on different platforms. If the two simulation systems have a similar interface structure – as is often the case with commercial bridge simulators developed according to similar

principles, then the zone configuration configured for one system can be transferred to another with minimal changes. This makes the method more universal. Unlike approaches that rely on direct access to internal simulator data, it is not rigidly tied to a specific software system.

At the same time, it is important to understand that the display of F depends on the configuration of the interface. This is a practical limitation of the method. If the interface structure changes noticeably – for example, after updating the software, when the indicator panels move or change location – the zones must be defined again. However, this approach immediately poses a practical question: how can these visual elements be systematically interpreted and converted into quantitative parameters suitable for scientific analysis? By itself, comparing the state of the interface with navigation variables gives only a general conceptual idea. In order to turn this idea into a real research tool, it is necessary to define a specific work procedure. It should allow to find the desired interface zones, interpret their graphic content and combine the obtained values into consistent sets of time-varying data.

6.4 Method architecture and patent protection

The proposed method is based on a simple sequence of actions, which includes four main stages: image capture, determination of the desired zone on the screen, interpretation of parameters and formation of time series (**Fig. 6.2**).

Stage 1. Image capture. The image from the simulator is fixed at equal time intervals Δt. The capture frequency is selected to match the refresh rate of the interface simulator indicators. In the experimental part, the interval $\Delta t = 1$ second (that is, the frequency of 1 Hz) was used as the base value. At the same time, the very architecture of the method allows working with a higher frequency of – up to 10 Hz, if technical conditions allow it. The resulting frames are used as downstream processing outputs.

Stage 2. Localization of zones. Before starting work, the analyst performs a preliminary configuration of the system. For each navigation indicator to be tracked, the screen sets its own Z_i zone. It is determined using normalized coordinates in the screen space. Once configured, the configuration is saved and used in all subsequent sessions. The setup process itself can be considered semi-automatic. First, the researcher manually marks the required areas on the screen, and then the system automatically applies this markup during further data processing.

Stage 3. Parameter interpretation. After the zones are defined, the system analyzes their content and converts it into numerical values. The method of processing depends on the type of indicator.

If the indicator shows the numbers – such as heading (HDG), speed relative to ground (SOG) or screw rotation (RPM) – optical character recognition is used. It simply reads the numbers displayed on the screen.

If the indicator has the form of a scale or dial – for example, the rate of rotation (ROT) indicator or the angle of translation of the steering wheel – geometric analysis is used. In this case, the system determines the position of the arrow or marker on the scale and then calculates the corresponding value.

Stage 4. Formation of time series. The parameter values obtained at each time point t_k are combined into one set. It can be written as a vector

$$P(t_k)=\{p_1(t_k),p_2(t_k),\ldots,p_n(t_k)\}.$$

Each element of this vector corresponds to a specific navigation or control parameter that is read from the simulator interface. Such sets are formed for each moment in time. If to place them one after the other, a multidimensional time series is formed. It shows how the state of the vessel has changed throughout the simulation.

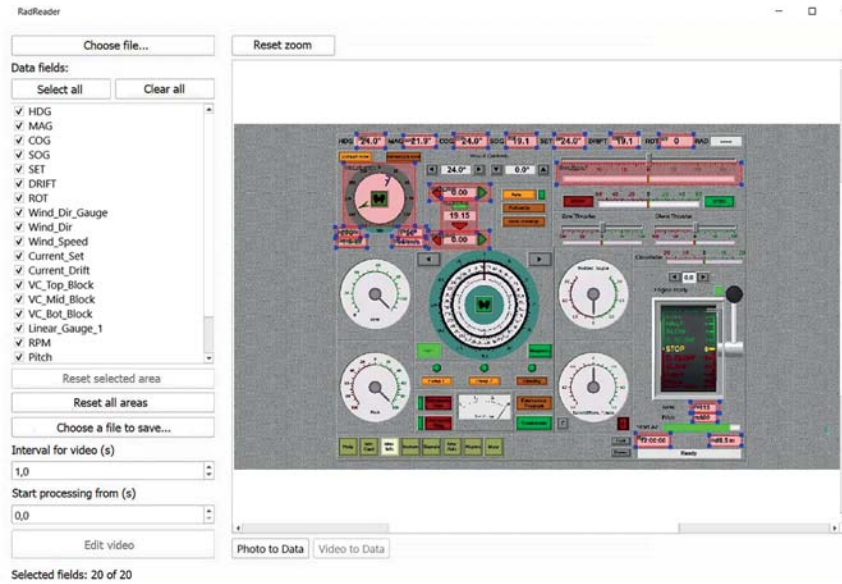


Fig. 6.2 Localization of information zones for reconstructing navigation parameters

Table 6.1 shows the main parameters that were obtained during the experiments. For each of them, it is indicated exactly how it is displayed on the simulator interface and how it is used in further analysis.

Table 6.1 Parameters obtained by the method

Parameter	Symbol	Display Type	Analytical role
Course	HDG	Numeric/Dial	Main trajectory parameter
Heading over ground	COG	Numeric	Indicator of ecological drift
Speed over ground	SOG	Numeric	Kinematic state variable
Turn rate	ROT	Numeric/Dash	Indicator of maneuver dynamics
Propeller revolutions	RPM	Numeric/Dial	Propellant state variable
Rudder angle	Rudder	Dial	Controlling input variable
Wind direction	Wind Dir	Numeric/Gauges	Environmental disturbances
Wind speed	Wind Spd	Numeric	External influences
Water depth	Depth	Numeric	Bathymetric context

One parameter value shows almost nothing by itself. It is much more important to see how these values change during the maneuver. Forming time series – is not just about combining data. It allows to move from individual measurements to the analysis of how the system behaves over time. It is this kind of sequence that helps to detect reaction delays, oscillations, stabilization processes and the relationship between the control actions and the ship's reaction.

Therefore, it is very important that all parameters are linked to the same timeline. For example, changing the course, shifting the steering wheel and changing the engine speed must correspond to the same moment in time. Only in such a case can the causal relationships between the control actions and the ship's response be correctly analyzed. In the proposed method, synchronization is ensured by the fact that all parameters are read from the same image frame. Due to this, all elements of the vector $P(t_k)$ correspond to the same moment in time during the simulation.

At the same time, the discrete nature of time series creates certain limitations. Events that occur within the same time interval Δt may not be fixed separately. However, for maneuvering analysis, this is usually not a significant problem, since such processes unfold quite slowly and their characteristic time is measured in tens of seconds. Therefore, a sampling rate of 1 Hz is sufficient in most cases.

If it is necessary to investigate fast transients in more detail, the method allows the use of a higher sampling rate. In practice, its maximum value is limited by the

speed of updating the simulator interface and the computing capabilities of the image processing system.

Data recovery from the visual display inevitably introduces some error compared to the reference values stored in the internal simulator model. Unlike direct access to telemetry, the proposed approach recovers parameter values through graphical interpretation, and accuracy depends on display resolution, indicator format, and graphical display stability.

Let $p_i^*(t)$ denote the reconstructed value of parameter I at time t, and $p_i(t)$ – its reference value within the simulator model. The instant recovery error is defined as follows

$$\varepsilon_i(t) = p_i^*(t) - p_i(t).$$

For the estimation over a given time interval, the mean absolute error (MAE) is used as the main indicator of accuracy

$$MAE_i = \frac{1}{N}\sum_k \left| p_i^*(t_k) - p_i(t_k) \right|,$$

where N – the number of discrete observation steps in the selected interval. This indicator is insensitive to the sign, which is suitable for estimating errors in the context of navigation, where both positive and negative deviations are equally significant.

The method described in this chapter is currently undergoing a patenting procedure. It was followed by two applications to the Ukrainian National Office for Intellectual Property and Innovation (UANIPIO):

– invention application No. a202600852;

– utility model application No. u202600851.

Both applications relate to the same technical solution. It is a method of automatically obtaining the navigation and control parameters of the vessel by analyzing the image of the navigation simulator interface, without referring to its internal telemetry data. Since the patent examination procedure has not yet been completed, the text uses the standard wording: "patent application filed, examination continues". Before possible commercial use of the method, it is recommended to clarify the current status of both applications on the official register of UANIPIO. The fact that two applications were submitted is related to the nature of the decision. On the one hand, the proposed approach contains new conceptual elements, so it is perceived as an invention. On the other hand, the method has quite obvious practical applicability, so it can be designed as a useful model.

6.5 Experimental verification and results

The method was validated on a fully functional bridge navigation simulator. Its interface shows a complete set of navigation and technical parameters that reflect the state of the vessel during simulation.

At the top of the interface are the main navigation indicators: heading (HDG), magnetic heading (MAG), heading relative to the ground (COG), speed (SOG), drift parameters and turning speed (ROT).

In the central part there is a circular course indicator that shows the orientation of the vessel in space.

At the bottom of the interface there are indicators of the speed of rotation of the propeller shaft (rpm), the angle of translation of the steering wheel, the speed of rotation and the engine controls.

The experimental program included several maneuvering scenarios. In these scenarios, the ship's course changed according to a given plan, while the engine operating parameters changed.

For each scenario, the parameter values recovered by the proposed method were compared with the reference values. These reference data were read directly from the simulator interface under controlled conditions. Such a comparison made it possible to calculate the indicators of the average absolute error.

Before starting the experiments, the zones corresponding to the different indicators were adjusted for the simulator interface.

For all the parameters shown in **Table 6.1**, as well as for some additional environmental variables displayed on the screen, separate observation areas were defined.

This installation was performed once. After that, the configuration was automatically used in all subsequent experimental sessions.

The coordinates of the zones were kept in a normalized form. Due to this, their position did not depend on the resolution of the screen. In tests where the resolution changed between different sessions, the configuration of the zones did not have to be adjusted again. This confirms the practical efficiency of the coordinate normalization approach.

Table 6.2 summarizes the results of the assessment of the accuracy of the main navigation parameters obtained during a series of experimental tests.

A course recovery error not exceeding 1° corresponds to the level of accuracy usually required when studying the maneuverability of a vessel and analyzing its trajectory. The error of determining the engine speed is less than 2% of the nominal value. This shows that the thrust state of the engine is reconstructed quite reliably over the entire range of modes used in the experimental scenarios.

Table 6.2 Reconstruction accuracy of basic navigation parameters

Parameter	MAE	Rating
Heading (HDG)	≤ 1.0°	Sufficient for maneuver analysis
Propeller RPM	< 2% nominal	Sufficient for motor setup research
Rotation Rate (ROT)	< 0.5 deg/min	Suitable for transient analysis
Speed Over Ground (SOG)	< 0.2 kt	Suitable for trajectory research

Fig. 6.3 is a smoothed time series of vessel heading (HDG) and yaw rate (ROT) during the active phase of the maneuver. The timeline covers approximately 700 seconds and includes the stage of starting the maneuver, the turn itself and further stabilization of the movement.

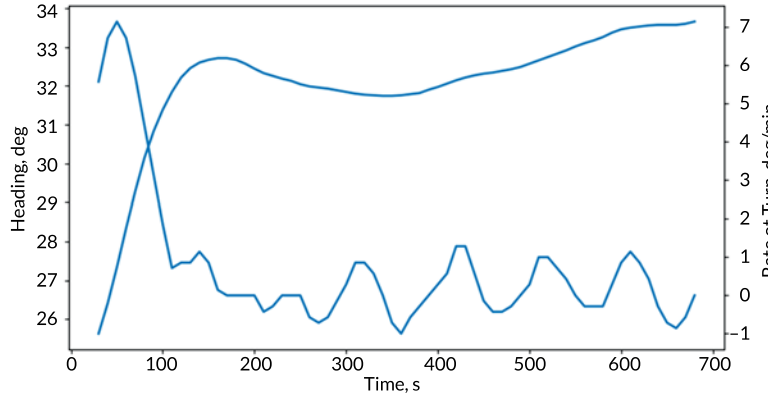


Fig. 6.3 Time series of vessel heading (HDG) and angular rate of rotation (ROT) on the maneuver interval
Source: Imagined with AI

The turn rate (ROT) time series shows a well-defined maximum that appears before a noticeable change in course. This corresponds to the real dynamics of the ship's movement: first, the angular speed of the turn increases, and then the course itself begins to change rapidly. When the ROT peak occurs earlier than the maximum rate of change of course dHDG/dt is a characteristic feature of a real maneuver and at the same time is an indirect check of the correctness of the performed data reconstruction.

There are no noticeable phase shifts or artificial fluctuations in the reconstructed time series, and this is especially important, because if such distortions occurred, it could lead to a misinterpretation of cause-and-effect relationships. For example, it might seem that the change of course occurs before the controlling influence,

although in reality this is not the case. And it is possible to see that the reconstruction method does not make such mistakes.

The time series of propeller revolutions (**Fig. 6.4**) at the same interval shows a gradual decrease in rotation frequency, which corresponds to the controlled transition of the vessel to a lower mode of movement. The synchronicity of the change of revolutions with the dynamics of the course and the angle of translation of the steering wheel allows analyzing the maneuver simultaneously from a kinematic and energy point of view. In other words, it becomes possible to consider the relationship between the mode of propulsion and the vessel's response to maneuver within a single coherent data set.

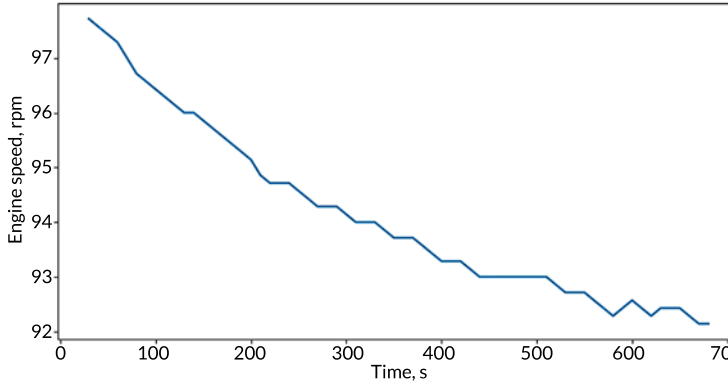


Fig. 6.4 Time series of engine speed (RPM) during the maneuver interval
Source: Imagined with AI

As a result of the experiments, it became clear that the proposed method allows to reconstruct navigation parameters quite stably and with reasonable accuracy, using only the image of the simulator interface. It can be applied to the analysis of maneuvering dynamics, the study of transient modes of motion and the comparison of different scenarios. At the same time, no connection to the internal software architecture of the simulator is required.

6.6 Conclusions

In this project, we took what is essentially a very simple idea and tested it in practice: what if, instead of digging into the simulator's inner workings, we simply

read what it already displays on the screen? To do this, we used image recognition. We stopped viewing the interface merely as a picture for the navigator and began to see it as a reflection of the model's internal state. Course, rudder angle, RPM – all of this is already right before our eyes. We just need to interpret it correctly. As it turned out, the approach works.

The main motion parameters are reconstructed quite reliably even without access to telemetry. The time series does not fall apart, events occur in the correct order, and the ship's behavior can be tracked dynamically. According to the experimental results, the sampling rate reached approximately 10 Hz, and the error remained within about 1° for the heading and up to 2% for the propeller RPM. For practical tasks, such as maneuver analysis or trajectory estimation, this is more than sufficient.

If to look at the bigger picture, the difference from typical approaches is quite noticeable here. Usually, a researcher connects to the simulator's internal mechanisms and retrieves logs, telemetry, or service channels. This is convenient, but there's a catch. Everything is tightly tied to a specific system. A different simulator means a new integration. Here, the situation is different. We don't connect to anything; instead, we work with what's already been exposed externally. Therefore, the method is easier to port between platforms. Essentially, it's not the tool that changes, but only the interface.

Of course, there are limitations. It all depends on how stable the interface is. However, if the layout of the elements changes, the system must be reconfigured. There is also a dependency on image quality, as a lack of sharpness or low resolution can affect accuracy, although this is not usually a critical issue under real-world conditions.

An important point is that this approach does not replace traditional methods of data collection. In other words, direct access to telemetry remains the simplest method, but if such access is unavailable or restricted, the option of screen reading becomes a perfectly viable alternative. This makes it possible to capture a consistent data set without interfering with the system software.

Looking at the bigger picture, the idea itself is interesting. Image recognition is usually associated with monitoring and diagnostics, but can also be used in experimental studies. In ship simulators, the interface is generally viewed as something for the human user, although it is actually data-level and quite extensive.

And this opens up scope for development. For example, it is possible to remove the manual configuration and train the system to find the necessary elements on the screen itself; it is possible to increase the processing frequency.

Overall, the main result of this work is that navigation data does not necessarily have to be taken from inside the simulator. It can be reconstructed from what is already visible. This makes experimental work simpler, more flexible, and less dependent on a specific system.

Conflict of interest

The authors declare that there is no conflict of interest in relation to this paper, as well as the published research results, including the financial aspects of conducting the research, obtaining and using its results, as well as any non-financial personal relationships.

Financing

The study was performed without financial support.

Use of artificial intelligence statement

The authors have used artificial intelligence technologies (ChatGPT, OpenAI) for language refinement and structuring of the manuscript. The authors bear full responsibility for the final manuscript. Generative AI tools are not credited and are not responsible for the final results.

Some figures in this manuscript were generated using artificial intelligence and are labeled accordingly.

Authors' contributions

Yevgeniy Kalinichenko: Conceptualization, Methodology, Supervision, Writing – review and editing.

Oleksandr Koliesnik: Investigation, Data Curation, Writing – original draft.

Oleg Safyan: Formal Analysis, Validation, Writing – review and editing.

References

1. Munim, Z. H., Krabbel, H., Haavardtun, P., Kim, T.-E., Bustgaard, M., Thorvaldsen, H. (2023). Scenario Design, Data Measurement, and Analysis Approaches in Maritime Simulator Training: A Systematic Review. Methodologies and Intelligent Systems for Technology Enhanced Learning, Workshops – 13th International Conference. Springer, 39–47. https://doi.org/10.1007/978-3-031-42134-1_4

2. Sellberg, C. (2016). Simulators in bridge operations training and assessment: a systematic review and qualitative synthesis. WMU Journal of Maritime Affairs, 16 (2), 247–263. https://doi.org/10.1007/s13437-016-0114-8
3. Tusher, H. M., Munim, Z. H., Nazir, S. (2023). An evaluation of maritime simulators from technical, instructional, and organizational perspectives: a hybrid multi-criteria decision-making approach. WMU Journal of Maritime Affairs, 23 (2), 165–194. https://doi.org/10.1007/s13437-023-00318-1
4. Hetherington, C., Flin, R., Mearns, K. (2006). Safety in shipping: The human element. Journal of Safety Research, 37 (4), 401–411. https://doi.org/10.1016/j.jsr.2006.04.007
5. Saeed, F., Bury, A., Bonsall, S., Riahi, R. (2018). A proposed Evidential Reasoning (ER) Methodology for Quantitative Assessment of Non-Technical Skills (NTS) Amongst Merchant Navy Deck Officers in a Ship's Bridge Simulator Environment. TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation, 12 (3), 597–608. https://doi.org/10.12716/1001.12.03.20
6. Atik, O. (2019). Eye tracking for assessment of situational awareness in bridge resource management training. Journal of Eye Movement Research, 12 (3). https://doi.org/10.16910/jemr.12.3.7
7. Ronca, V., Uflaz, E., Turan, O., Bantan, H., MacKinnon, S. N., Lommi, A. et al. (2023). Neurophysiological Assessment of An Innovative Maritime Safety System in Terms of Ship Operators' Mental Workload, Stress, and Attention in the Full Mission Bridge Simulator. Brain Sciences, 13 (9), 1319. https://doi.org/10.3390/brainsci13091319
8. Frette, V., Kleppe, G., Christensen, K., Garrido, P. L., Marro, J., de los Santos, F. (2011). Analysis of ship maneuvering data from simulators. AIP Conference Proceedings, 1332, 276. https://doi.org/10.1063/1.3577668
9. Li, H., Zhao, N., Zhou, J., Chen, X., Wang, C. (2024). Uncertainty Analysis and Maneuver Simulation of Standard Ship Model. Journal of Marine Science and Engineering, 12 (7), 1230. https://doi.org/10.3390/jmse12071230
10. Pires da Silva, P., Sutulo, S., Guedes Soares, C. (2023). Sensitivity Analysis of Ship Manoeuvring Mathematical Models. Journal of Marine Science and Engineering, 11 (2), 416. https://doi.org/10.3390/jmse11020416
11. He, H., Lataire, E., Zwijnsvoorde, T. V., Delefortrie, G. (2023). A Ship Manoeuvring Desktop Simulator for Developing and Validating Automatic Control Algorithms. TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation, 17 (3), 607–616. https://doi.org/10.12716/1001.17.03.12

12. Yasukawa, H., Yoshimura, Y. (2015). Introduction of MMG standard method for ship maneuvering predictions. Journal of Marine Science and Technology, 20 (1), 37–52. https://doi.org/10.1007/s00773-014-0293-y
13. Lee, J., Nam, B., Lee, J.-H., Kim, Y. (2021). Development of Enhanced Two-Time-Scale Model for Simulation of Ship Maneuvering in Ocean Waves. Journal of Marine Science and Engineering, 9 (7), 700. https://doi.org/10.3390/jmse9070700
14. Kim, D., Tezdogan, T., Incecik, A. (2022). A high-fidelity CFD-based model for the prediction of ship manoeuvrability in currents. Ocean Engineering, 256, 111492. https://doi.org/10.1016/j.oceaneng.2022.111492
15. Smith, R. (2007). An Overview of the Tesseract OCR Engine. Ninth International Conference on Document Analysis and Recognition (ICDAR 2007), 2, 629–633. https://doi.org/10.1109/icdar.2007.4376991
16. Szeliski, R. (2011). Computer Vision: Algorithms and applications. Texts in Computer Science. London: Springer. https://doi.org/10.1007/978-1-84882-935-0
17. Wuest, T., Weimer, D., Irgens, C., Thoben, K.-D. (2016). Machine learning in manufacturing: advantages, challenges, and applications. Production & Manufacturing Research, 4 (1), 23–45. https://doi.org/10.1080/21693277.2016.1192517
18. Sami Ur Rehman, M., Shafiq, M. T., Ullah, F. (2022). Automated Computer Vision-Based Construction Progress Monitoring: A Systematic Review. Buildings, 12 (7), 1037. https://doi.org/10.3390/buildings12071037

CHAPTER 7

BPMN as a tool for adaptive support of pattern recognition and digital diagnostic results in remediation and post-crisis recovery

Bohdan Cherniavskyi
Oksana Drozd
Anatolii Nadtochyi
Viktor Nadtochii
Maksym Matviienko
Ivan Kalinichenko

Abstract

In the monographic study, a comprehensive analysis of the potential of BPMN 2.0 (Business Process Model and Notation 2.0) as a tool for adaptive support of pattern recognition results and digital diagnostics in the processes of remediation and post-crisis recovery of objects and territories is presented. The relevance of the topic is due to the fact that modern systems of monitoring, remote sensing, computer vision, and intelligent analytics are capable of promptly identifying the consequences of emergency events of natural and technogenic nature, namely: damage to infrastructure objects, contamination of territories, and other crisis consequences. However, at the same time, a managerial and organizational-technological gap often remains between the digital detection of a problem and the actual implementation of reconstruction and remediation measures. In this context, BPMN (Business Process Model and Notation) is proposed to be considered as a process tool that is capable of ensuring the formalization, routing, as well as coordination of further managerial actions based on the results of digital diagnostics. In the monograph, the concept of adaptive support is substantiated, within the framework of which pattern recognition results are interpreted as a basis for launching a managed process of identification, verification, refinement, expert evaluation, selection of an appropriate response protocol, as well as resource and logistical support with subsequent verification of the results of the implemented remediation measures. The authors pay particular attention to the process of handling information with varying degrees of uncertainty, which is associated, among other things, with the probabilistic nature

 DOI: 10.21303/978-9908-845-05-0.ch7

of AI (Artificial Intelligence) outputs, including the use of confidence thresholds, escalation mechanisms, as well as repeated data collection and feedback. The monograph substantiates that the application of BPMN (Business Process Model and Notation) makes it possible to link the analytical layer of digital diagnostics with the operational layer of remediation into a unified process logic, thereby increasing the level of transparency, traceability, as well as coordination of actions of key agents and stakeholders.

Keywords

Business process model and notation (BPMN), pattern recognition, digital diagnostics, remediation, post-crisis recovery, computer vision (CV), decision model and notation (DMN), case management model and notation (CMMN), human-in-the-loop (HITL), digital passport of territory.

7.1 Introduction

It should be noted that in the first quarter of the XXI century post-crisis recovery of territories is increasingly determined by the ability of management systems to operationally transform fragmented data into coordinated managerial actions. This is due to the fact that modern risks have a systemic and cascading nature, simultaneously affecting infrastructure, the natural environment, supply chains, institutional mechanisms, and social resilience.

For this reason, the tasks of remediation and post-crisis recovery can no longer be considered as a set of disparate organizational and technical measures. On the contrary, they imply the integration of diagnostic, analytical, organizational, and execution contours into a unified logic of developing and implementing effective managerial decisions. Thus, in the published documents of UNDRR (United Nations Office for Disaster Risk Reduction), in particular in the Global Assessment Report, it is emphasized that modern risks propagate across sectors and territories, and effective loss reduction requires the evolutionary development of management systems and a transition to more coordinated models of coordination [1]. In this context, the search for such instruments that are capable of ensuring a controlled transition from the digital identification of damage to the actual implementation of recovery and remediation measures acquires particular significance.

Additional relevance to the topic is provided by the rapid development of remote sensing, satellite monitoring, computer vision, and intelligent analytics. Modern operational observation systems already make it possible to map natural and technogenic threats, identify damage to buildings and other objects, assess the condition

of infrastructure, and record crisis changes across large territories. NASA (National Aeronautics and Space Administration) emphasizes that Earth satellite data are used to map natural hazards and reduce damage from floods, fires, and other disasters, while ARSET (Applied Remote Sensing Training) programs demonstrate their practical application for disaster risk assessment, resilience provision, and humanitarian monitoring. In turn, CEMS (Copernicus Emergency Management Service) already provides cartographic support for the management of natural and anthropogenic emergencies, humanitarian crises, as well as tasks of post-conflict recovery and reconstruction planning. This indicates that the technological base for the operational acquisition of digital information on the nature of damage and the spatial distribution of consequences has already been formed. At the same time, the availability of diagnostic data in itself does not yet guarantee their effective transformation into a sequence of managerial and logistical actions [2–5].

Practice shows that one of the key barriers in the management of crises and emergencies remains the gap between the stages of damage detection and the implementation of managerial decisions for consequence elimination and remediation. Even in the presence of high-quality results of pattern recognition, a number of unresolved issues still remain, related to their verification, prioritization, alignment with regulatory protocols, allocation of actually available resources, involvement of experts in selecting the optimal scenario, and re-validation of results after the completion of the remediation work complex.

The situation is further complicated by the fact that the outputs of AI (Artificial Intelligence) models often retain a probabilistic nature and are quite sensitive to domain shifts, the type of disaster, the specific features of the territorial distribution of damaged objects, and the characteristics of the input data. Thus, studies based on xBD/xView2 confirm that automated damage assessment of buildings using satellite imagery is a significant direction for disaster response and recovery, however, it requires unified damage scales and more precise interpretation of results. The analysis of recent scientific works also demonstrates the fact that the generalization capability of models to new challenges and threats of a hybrid nature still remains limited, whereas adaptive methods yield better results when the disaster context is taken into account. Consequently, for remediation tasks, not only the recognition algorithm itself is of fundamental importance, but also the process mechanism of supporting its result under conditions of uncertainty [6–10].

It is precisely in this, in the opinion of the authors of the monograph, that the study and practical significance of BPMN (Business Process Model and Notation) lies. According to the specification of Object Management Group (Object Management Group), BPMN is a formal standard for business process modeling that

combines clarity for stakeholders with a level of precision sufficient for translating diagrams into software process components.

This makes BPMN potentially suitable for constructing an end-to-end logic of coordination in complex interagency and multivector remediation processes, especially in situations of elimination of emergencies of natural and technogenic nature, where it is necessary to manage not only the sequence of actions, but also exceptions, multivariant scenarios, events, the alignment of various organizational, economic, technical and other decisions, as well as the necessity of direct participation of authorized persons within the management loop. Studies on process organization in crisis situations indicate the BPMN advantages and the expediency of its combination with CMMN (Case Management Model and Notation) and DMN (Decision Model and Notation) for the purpose of supporting more flexible knowledge-intensive scenarios. Therefore, an in-depth study of BPMN not as a recognition tool, but as a tool for adaptive support of digital diagnostics results appears logical and methodologically justified. The practical value of such an approach consists in overcoming the fragmentation of digital crisis management contours through process logic that allows for data refinement, repeated data collection, expert verification, and context-dependent case routing. No less important additional advantage remains the possibility of its validation based on open data from Copernicus Emergency Management Service, xBD/xView2, and SpaceNet 8 [11–14].

The objective of the study is to substantiate BPMN as a tool for adaptive support of pattern recognition results and digital diagnostics in remediation and post-crisis recovery processes. To achieve the stated objective, it is necessary to solve three interrelated tasks. Firstly, to reveal the specifics of the managerial gap between the digital identification of consequences of emergencies of various nature and scale and the actual implementation of recovery measures, as well as to determine the place of Pattern Recognition results and digital diagnostics within the overall architecture of managerial decision-making in the remediation of affected objects and territories. Secondly, to substantiate the possibilities of using BPMN for the purpose of formalizing routes of verification, escalation, coordination, and control under conditions of uncertainty of AI outputs. Thirdly, to identify the practical applicability of the proposed approach for constructing controllable digital response contours and to outline the possibilities of its verification based on the composition of open benchmark datasets, crisis geodata, and process metrics for evaluating orchestration.

The solution of these tasks will make it possible to clarify the theoretical foundations of process orchestration of digital diagnostics and, at the same time, to enhance the practical applicability of BPMN in complex scenarios of remediation and post-crisis recovery.

7.2 State of the study of the problem and literature review

First of all, it should be emphasized that the state of the study of the problem under investigation is characterized by a fairly high degree of fragmentation. Thus, the scientific literature widely presents studies on process modeling, digital diagnostics of the consequences of emergency events of a crisis nature, remote sensing, and automated assessment of incurred damage; however, their integration into a unified logic of adaptive support of recognition results for the tasks of remediation and post-crisis recovery remains rather limited. It should be noted that the theoretical and methodological layer directly related to BPMN as a standard for business process modeling has been studied and developed in sufficient depth. The OMG specification establishes BPMN as a generally accepted notation oriented simultaneously toward comprehensibility for all key agents and stakeholders and, at the same time, toward sufficient formal precision for translation into software process components. According to the conviction of the authors of the monograph, it is precisely this dual nature of BPMN, as well as the combination of visual transparency and formal certainty, that has determined its wide dissemination in the tasks of regulation, automation, and coordination of processes in various subject areas [8, 11].

At the same time, studies show that classical BPMN is most effective in cases where the process can be represented as a relatively structured flow of actions. Under conditions of emergencies and weakly predictable scenarios of event unfolding, procedural logic alone often proves insufficient, which has led to increased interest in the combined use of BPMN, CMMN, and DMN. Within such a combination, BPMN makes it possible to ensure flow orchestration, CMMN provides case-oriented flexibility, and DMN supports managerial decision-making. As noted in their scientific work by Niemz, Gehrke, and Ruhland, it is precisely such an integrated architecture that is better adapted to the changing conditions of crisis situations than the isolated application of a single notation. A similar conclusion is also contained in studies on adaptive case management in emergency response, where it is emphasized that hybrid process models make it possible to adapt execution more rapidly to the development of an incident while simultaneously preserving controllability under conditions of increasing uncertainty [8].

Another significant body of the study is associated with the digital diagnostics of the consequences of crisis phenomena of natural, technogenic, and military nature based on remote sensing, computer vision, and Pattern Recognition methods. In recent years, this direction has advanced substantially due to the development of open datasets and benchmark platforms. Of particular importance is the xBD dataset, created as a large-scale open resource for change detection and building damage

assessment in the interests of humanitarian assistance and disaster recovery. The authors of xBD emphasize that after the occurrence of emergency events, rapid assessment of the actual condition and incurred damage, as well as logistical planning of evacuation, delivery, and distribution of resources, are critically important. In this context, satellite imagery in combination with computer vision is capable of significantly reducing the burden on physical verification through inspection. Additionally, SpaceNet 8 makes it possible to expand this approach by integrating tasks of recognition of buildings, roads, and other objects, thereby extending the horizon of functional capabilities not only to damage detection but also to the analysis of post-crisis response and the successful remediation of affected territories [11].

In contemporary science, the capabilities of tools for the primary digital identification of damage and environmental changes associated with the occurrence of emergency events have been studied in considerable depth. Thus, operational platforms such as CEMS confirm that, in practice, a mature infrastructure for obtaining open geospatial data for emergency response, risk, and recovery mapping has already been formed. In particular, Copernicus Emergency Management Service explicitly indicates that recovery mapping is used for the purposes of supporting recovery process management, reconstruction planning, agricultural and natural area recovery, as well as the formation of post-crisis resilience [15]. The above makes it possible to conclude that the layer of digital diagnostics and spatial documentation of damage is, in general, significantly more developed than the layer of subsequent process support of the obtained results [16, 17].

A separate direction of scientific literature is associated with the study of the problems of applying remote sensing and spatial methods for the assessment of pollution and environmental monitoring, which are directly related to the processes of elimination of contamination of various nature and the implementation of remediation measures. Publications of USGS (United States Geological Survey) and contemporary review studies show that remote sensing can be used for the identification of residual contamination, the assessment of vegetative stress as a proxy indicator of pollution, the mapping of hotspot zones, and the support of environmental decision-making in affected territories. Recent works on monitoring soil contamination and technogenic impacts confirm that geodata and machine learning methods significantly expand the horizon of environmental diagnostics capabilities; however, in most cases they still remain tools of observation and analytics rather than elements of a formalized process for translating results into specific managerial actions. Consequently, within the environmental and remediation contour, the same pattern manifests as in the field of disaster damage assessment: diagnostics develops faster than the process orchestration of its results [18–22].

In addition, one of the significant directions of in-depth scientific research is constituted by publications on process mining and data-driven approaches in the context of process execution analysis. The BPI Challenge and the open event logs associated with it have effectively formed a benchmark environment for testing methods of process mining, predictive monitoring, and bottleneck analysis. According to the view of the authors of the monograph, the significance of this direction within the raised problem lies in the existence of an already established apparatus of metrics and approaches for evaluating the actual behavior of processes, including delays, deviations, repeated actions, and bottlenecks in the implementation of operations. At the same time, a comprehensive survey of publications on this topic has revealed the fact that available benchmark logs, as a rule, relate to banking, administrative, medical, and other organizational processes, rather than to the elimination of the consequences of emergency events and subsequent remediation, and not to the support of AI results in crisis systems. Consequently, the methodological toolkit for evaluating orchestration already exists, whereas the domain-specific event base for remediation remains practically undeveloped [23, 24].

An important layer of contemporary research is constituted by publications reflecting issues related to HITL (Human-in-the-Loop) and human oversight in high-risk AI systems. Thus, studies emphasize that in high-stakes domains a human must retain control over critical decisions, while AI acts as a means of supporting them. Within the logic of the present monographic study, this is of fundamental importance, since the results of pattern recognition in crisis diagnostics often have a probabilistic nature and require verification, refinement, and contextual interpretation. At the same time, the literature on HITL predominantly develops within the framework of AI governance, explanation, reliability, and decision support, rather than within the framework of formalized BPMN orchestration of the full incident life cycle – from recognition to the completion of the remediation case [25–28].

Thus, a comprehensive analysis of literary sources makes it possible to draw several generalizing conclusions. In contemporary science, such directions as process modeling based on BPMN, hybrid crisis management models BPMN/CMMN/DMN, digital damage assessment based on satellite data, open crisis geodata, and process mining methods have already been sufficiently well developed. To a significantly lesser extent, the question of how the probabilistic result of Pattern Recognition and digital diagnostics can be transformed into an adaptively supported, verifiable, and controllable process of remediation and post-crisis recovery has been studied. At least four bottlenecks remain that require in-depth investigation: the absence of

a domain-specific orchestration model of recognition results for remediation, insufficient elaboration of process handling of uncertainty of AI outputs, weak integration of HITL into executable process contours, and the absence of a direct open benchmark for the verification of such hybrid systems. It is precisely the combination of these gaps that forms a research niche for further substantiation of BPMN as a tool for adaptive support of pattern recognition results and digital diagnostics in remediation and post-crisis recovery [8, 11, 17, 25].

The results of the problem-oriented analytical cross-section are presented below in **Table 7.1**.

Table 7.1 Degree of study of the problem by blocks across research directions

Research Direction Block	Representative Sources	What Has Been Studied to Date	Bottlenecks / What Has Been Insufficiently Studied	Theoretical Maturity (0–5)	Practical Development (0–5)
1	2	3	4	5	6
BPMN as a process modeling standard	OMG BPMN 2.0 (2010–2013)	Formal notation, semantics of elements, modeling and automation of processes	The role of BPMN in supporting probabilistic AI outputs is insufficiently disclosed	5	5
BPMN in crisis and weakly predictable scenarios	Niemz, Gehrke, Ruhland (2021)	Application of BPMN for crisis situations and coordination of actions	Insufficient number of cases for remediation and post-crisis recovery	3	2
Integration of BPMN + CMMN + DMN for adaptive management	Niemz et al. (2021) and related works	Hybrid management of structured and unstructured scenarios, decision support	Insufficient adaptation to supporting Pattern Recognition results	4	2
Pattern Recognition / Computer Vision for damage assessment	Gupta et al., xBD (2019)	Damage recognition, change detection, building damage assessment	Recognition results are not translated into formalized process support	4	4
Open benchmark datasets for crisis visual diagnostics	xBD, xFBD, SpaceNet 8	Benchmark evaluation of damage detection, changes, infrastructure disruptions	No benchmark linkage with decision orchestration and closure of remediation cases	4	4

Continuation of Table 7.1

1	2	3	4	5	6
Open crisis geodata and recovery mapping	Copernicus EMS	Emergency, preparedness, and recovery mapping, damage assessment products	Geodata are weakly integrated into executable BPMN contours of adaptive support	4	4
Process mining and benchmark culture of event logs	BPI Challenge, Process Mining Task Force	Methods of execution analysis, bottlenecks, rework, deviations, process performance	Absence of open remediation / case-orchestration event logs	5	4
HITL in high-stakes AI	Contemporary HITL studies	Human control, expert verification, oversight	Weak formalization of HITL integration into BPMN remediation contours	3	2
Integration of AI diagnostics and process orchestration	Related works on workflow and crisis management	Partial linkage of analytics, workflow, and decision support	No stable end-to-end model of adaptive orchestration	2	1
Adaptive support of pattern recognition results and digital diagnostics in remediation and post-crisis recovery	Direct scientific sources not identified	Only related conceptual and applied foundations are available	Absence of an integral model, mature terminology, and direct benchmark	1	0–1

Note: Evaluation scale: 0 – practically not represented; 1 – isolated formulations; 2 – fragmentarily studied; 3 – moderately studied; 4 – well studied; 5 – comprehensively studied / standardized / supported by developed practice. Representative blocks of the table are based on the BPMN standard of the OMG, open disaster benchmarks xBD and SpaceNet 8, recovery mapping of the Copernicus Emergency Management Service, and the benchmark culture of process mining

7.3 Conceptual foundations of adaptive support of pattern recognition results by BPMN means

The conceptual idea of adaptive support of Pattern Recognition results by BPMN means should be presented based on the distinction between the very fact of digital recognition and the subsequent management based on these results. It refers to the fact that in crisis and post-crisis scenarios, pattern recognition forms a diagnostic conclusion regarding the nature and scale of damage, contamination, change of an

object and/or territory, as well as other anomalies; however, such a conclusion is not yet identical to a managerial decision. It is precisely for this reason that the result of Pattern Recognition should be considered not as the final point of analytics, but as an input to a subsequent process of verification, routing, and coordination of further managerial actions undertaken. In this sense, adaptive support means the organization of a controlled life cycle of the recognition result – from its initial fixation to the closure of the case after the implementation of remediation, reconstruction, and other recovery measures [11, 17, 29].

The basic logic of the proposed approach is grounded in the BPMN perception as a formal standard for process description that combines comprehensibility for all agents and stakeholders with semantic precision sufficient for software implementation. By virtue of this, BPMN may be considered as a process framework which, based on the results of digital diagnostics, makes it possible to structure an organized logic of subsequent managerial actions. The first conceptual foundation of the approach thus lies in the differentiation between the analytical layer and the orchestration layer; at the same time, BPMN does not substitute the AI model, but ensures the management of the subsequent remediation process within the decision-making system [29, 30].

The second foundation is related to the probabilistic nature of Pattern Recognition results. As demonstrated by publicly available benchmark datasets, in particular xBD, post-crisis damage assessment based on satellite data depends on the quality of input data, the disaster context, and the structure of damage classes. Consequently, the recognition result must be interpreted not only in terms of its content, but also in terms of the level of confidence, completeness, and contextual relevance, and the process support system must distinguish at least three situations: sufficient confidence for automatic transition to action, intermediate confidence – for additional verification, and insufficient confidence – for escalation, repeated data collection, or deferred decision-making [11].

The third conceptual foundation lies in the fact that adaptive support should not be reduced to a linear workflow as such. Thus, if the result of pattern recognition enters the system as a signal of a potential incident (for example, a destroyed infrastructure object, a contaminated territory, etc.), then in this case the subsequent logic must take into account not only the type of detected pattern, but also the context of its occurrence, key priorities, temporal and resource constraints, regulatory requirements, as well as the necessity of participation of specific actors (including experts, military personnel, and other specialists). Consequently, at the center of the model there should be not a rigid scenario, but a controlled case routing that provides for deviations, returns to previous stages and operations, possible branching options, as well as repeated verification cycles. Within this logic, BPMN acts as the

core of the structured part of orchestration, whereas adaptability is ensured through event-based logic, gateways, as well as subprocesses and tasks of specific actors of the remediation process and stakeholders with the integration of external rules. From this perspective, adaptive support is interpreted by the authors as a process mode in which the result of pattern recognition is not automatically accepted as a final basis for a specific managerial action, but passes through a contour of additional refinement, evaluation and forecasting, as well as selection of the further path and control of execution of a specific remediation scenario.

It should be noted that a key principle of the model proposed by the authors is the event-driven nature of interaction between the diagnostic and execution contours. Thus, within this logic, the Pattern Recognition result should enter the BPMN process not in the form of an abstract report, but as a structured event that reflects the type of pattern with geospatial reference, determination of the confidence level based on actual data sources, time of detection, and a preliminary assessment of the priority class. It is precisely this that makes it possible to automatically initiate the process of elimination of the consequences of an emergency event, trigger the corresponding subprocess, or embed a new case into an already operating management contour.

According to the authors of the monograph, the principle of multi-level data verification is no less important. Thus, in high-risk and post-crisis domains, it is unacceptable to build the entire management cycle on a single digital output without verifying its level of reliability. Therefore, in the proposed model, at least automatic verification according to a set of rules is provided, as well as the possibility of verification through the use of additional data sources and/or expert evaluation. In this regard, Human-in-the-Loop (HITL) occupies a special place in this logic as an institutionalized point of process escalation, namely: human participation acts as an embedded element of adaptive support of the remediation process management.

In addition to the above, no less significant is the principle of context dependency of the post-crisis management process. Thus, the same recognition result may require different responses depending on the type and scale of contamination, characteristics of the territory, accessibility of the area for the implementation of remediation measures, weather conditions, availability of laboratories, reagents, secondary risks of contamination or its scaling, as well as existing regulatory constraints. In this regard, the BPMN process must not only record the presence of identified patterns, but also correlate them with the parameters of the domain context. This determines the expediency of the architecture proposed by the authors, in which BPMN is responsible for the entire complex of remediation measures, while the decision layer or rule-based logic determines the conditions for route selection, threat level, and type of response protocol. For the domain of remediation and post-crisis recovery,

such logic is of particular importance, since the corresponding processes are multi-stage and, in a number of cases, rather long-term. Consequently, BPMN as a tool of adaptive support must ensure not a one-time switch to a specific managerial action, but the maintenance of the management case throughout several phases – from the initial assessment of contamination to confirmation of the effect of the implemented remediation measures complex.

From this follows the principle of a closed loop, which consists in the fact that adaptive support of the remediation management process must be completed by repeated acquisition of digital data, their comparison with the initial state of the affected object, and fixation of the achieved result. In generalized form, such a model includes six basic stages, namely: registration and structuring of the Pattern Recognition result; initial assessment of the level of its confidence and criticality; selection of the verification route; assignment of tasks and coordination of response/remediation actions; monitoring of the execution of the complex of works for elimination of the consequences of the emergency event; repeated digital verification and case closure.

As a result of a comprehensive analysis, the authors of the monograph have established that the availability of open benchmark datasets for recognition of damage incurred as a result of emergency events, geospatial recovery products, and process metrics for evaluating the execution of remediation operations makes it possible to consider this model as not only theoretically substantiated, but also potentially verifiable within a compositional research framework [8, 29, 31].

The scheme presented above reflects an architecture in which the result of pattern recognition is presented not as a final analytical product, but as an input to a controlled process of supporting the remediation of an affected object or territory. Thus, at the initial levels, data from heterogeneous sources are accumulated and integrated, after which the Pattern Recognition layer generates a primary digital output regarding the characteristics of damage, its parameters, associated risks, etc. This output is then transformed into a structured event containing the type of pattern, geospatial reference, determination of the confidence level, a list of data sources, as well as key priorities and constraints, which makes it possible to embed digital diagnostics into an executable management contour.

The central element of the model is the BPMN contour of adaptive support of the remediation process, within which case registration, assessment of the level of its criticality and data reliability, selection of the scenario/route of further actions, additional verification, or expert evaluation are carried out. As mentioned earlier, Human-in-the-Loop in the proposed architecture acts as an embedded component rather than an external addition, since the participation of experts (including military personnel, technical specialists, etc.) ensures confirmation, refinement, or rejection

of the obtained AI output. After that, the BPMN contour makes it possible to link the analytical data layer with practical actions for the elimination of the consequences of an emergency event, while ensuring the selection of an optimal remediation scenario with subsequent coordination of process participants and control of its execution.

The execution layer encompasses all participants of the remediation process who interpret the data of digital diagnostics, namely: contractors, regulators, military personnel, technical specialists, logistics operators, laboratory staff, expert groups, etc. Of fundamental importance is the feedback verification layer, which closes the architecture into a complete controlled cycle. Thus, after the implementation of the complex of remediation measures, repeated monitoring and comparison of the state of the affected object/territory according to the "before/after" scheme are carried out. At the same time, full case closure is permitted only after confirmation of the achieved restoration effect, which ensures the result-oriented nature of the model. Finally, the concluding level is the formation of a digital passport of the object/site, reporting, as well as the accumulation of lessons learned. In this context, process mining and process analytics create the basis for subsequent final evaluation of orchestration effectiveness and improvement of the remediation management system.

The conceptual architectural scheme is presented below in **Fig. 7.1**.

It should be emphasized that the Digital Passport of the affected object or territory is one of the central results of the architecture proposed by the authors, since it is precisely it that makes it possible to interlink digital diagnostics, remediation/recovery actions, the results of repeated data verification, institutional memory, and the reuse of the body of information from various sources upon the occurrence of new emergency events. Within the framework of this model, it is advisable to interpret it as an integral digital artifact of the state, the entire history of interventions, and the verified result for a specific spatial remediation unit. In other words, it is a specific structured, spatially referenced digital object in which both its initial and current state are recorded, the results of digital diagnostics, the entire history of interventions related to the elimination of the consequences of an emergency situation, information on the implemented remediation scenarios, the results of the conducted post-verification, as well as the current status of operational suitability, level of risk and/or restrictions. Thus, the Digital Passport ensures full digital traceability of all stages of management, namely: from the identified damage or contamination to the confirmed result of the implemented complex of remediation measures, accumulates institutional memory, supports intersystem exchange between GIS and the BPMN platform, as well as monitoring systems and analytical contours. At the same time, it also forms the basis for repeated monitoring, reassessment of the level of risk, and the initiation of a new remediation process cycle in the event of emergency situations.

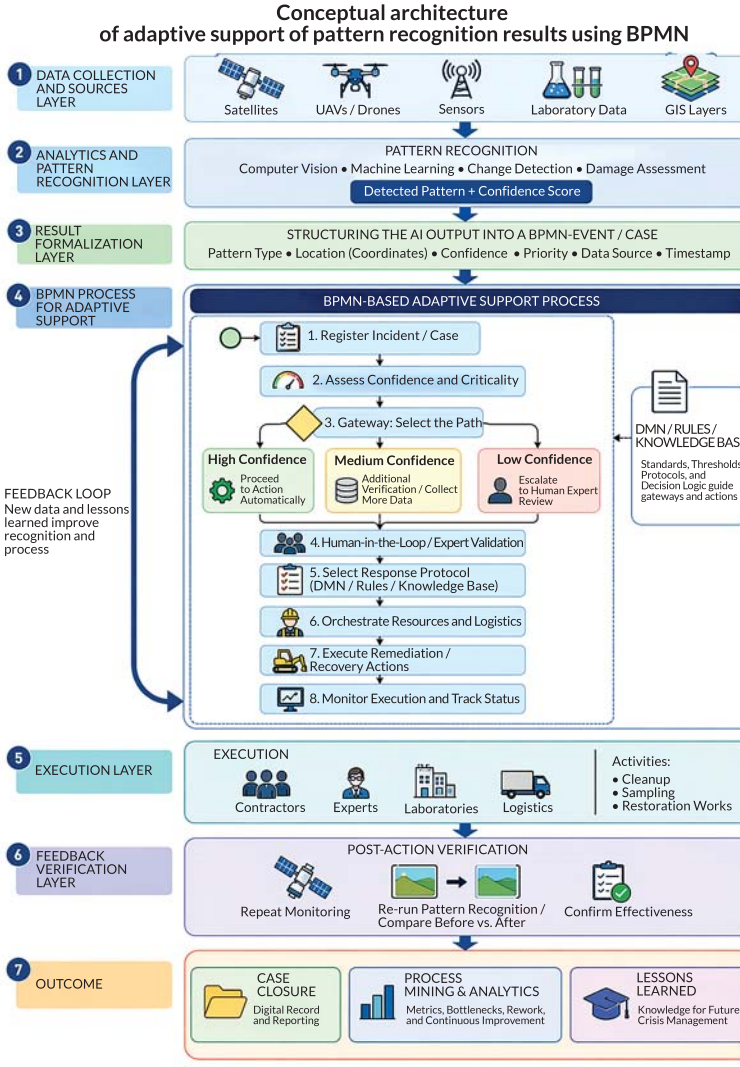


Fig. 7.1 Conceptual architecture of adaptive orchestration of pattern recognition results using BPMN
Note: images, photorealistic images, diagrams, drawings, figures that have been generated by artificial intelligence should be labeled "Imagined with AI".
Digital Passport of the territory as the final artifact of the BPMN contour of adaptive support of the remediation management process

From the perspective of the architecture of BPMN contours, the Digital Passport of the territory is interpreted by the authors as the final artifact of the closure stage of the remediation process of the affected object, which is formed after the completion of the full cycle of adaptive support. At a minimum, it should include the identifier of the object or affected territory, the type and nature of consequences, geometry and coordinate reference, the results of Pattern Recognition and digital diagnostics, the confidence level and verification status, information on the BPMN support route, as well as the list of implemented cleaning and recovery measures, the results of repeated verification, the current status of the remediation object, the presence of operational risk, as well as related digital artifacts. In the context of practical implementation, the basic format of its representation may be GeoJSON with an extended set of attributes, which it is advisable to supplement with structured JSON schemas for the purpose of integration with external services. For these purposes, it is also possible to use JSON-LD (JavaScript Object Notation for Linked Data) or RDF (Resource Description Framework). Unlike a one-time report, the Digital Passport should be considered as a dynamically updated digital object with its own life cycle – from creation and phased enrichment to final validation, reuse, and revision under new crises.

Thus, it acts not as a secondary appendix to the BPMN contour, but as its substantively significant outcome and, at the same time, as a mechanism for consolidating the verified state of the territory. The generalized vision of the structure of the Digital Passport of the territory is presented in **Table 7.2**.

The structure presented above demonstrates the fact that the Digital Passport of the territory should be considered as a final digital artifact that makes it possible to integrate spatial, diagnostic, process-related, and verification information about the affected object, including the site on which it is located. It records not only the initial incident, but also the full cycle of adaptive support – from case registration to confirmation of the result of the complex of implemented operations – while ensuring traceability of managerial decisions made, transparency of implemented measures, and accumulation of institutional memory. Due to the possibility of phased updating, the Digital Passport may be suitable not only for reporting, but also for reuse in the event of other crisis situations.

Summarizing all of the above, it may be argued that the most significant aspect, in the opinion of the authors, is the possibility of managing the level of uncertainty, namely: the presence of a confidence score means that the system can change the case route depending on the level of confidence, context, and criticality of the incident, rather than being limited to the binary scheme "recognized/not recognized". Therefore, in the case of an insufficient confidence level, the process must provide for additional verification, expert involvement, and/or repeated data collection. Such an approach makes it possible to bring the case from the recognition stage to the selection

of a protocol, the launch of the resource-use process, repeated verification, and formation of the Digital Passport of the site, which gives the idea of end-to-end orchestration both theoretical and practical validity. The next logical step of the study is the routing scheme by confidence level (**Fig. 7.2**).

Table 7.2 Structure of the digital passport of the territory

Passport Element	Characteristics / Composition of Information	Source of Formation	Stage of Update
Identification Block	Identifier of the passport and the site, type and functional purpose of the territory	BPMN-system, GIS, object registries, planning documentation	Creation of the passport; upon status change
Spatial Block	Coordinates, contour, geometry, and spatial localization of the site	GIS, satellite data, Copernicus Emergency Management Service, local geodata	Creation; refinement during repeated verification
Crisis Impact Block	Type of incident, date of detection, source of the primary signal	Event layer, AI/ML system, UAVs (Unmanned Aerial Vehicles), sensors, external reports	Initial case registration
Diagnostic Block	Type of identified pattern, class of damage/contamination, results of digital diagnostics	CV pipeline, AI/ML-model, analytical layer	Initial registration; repeated diagnostics
Confidence and Verification Block	Confidence level of the output, confirmation status, expert conclusions	CV pipeline, BPMN/DMN, expert verification	Verification; repeated assessment
Prioritization and Routing Block	Case criticality, selected BPMN support branch, assigned response protocol	DMN, rule engine, BPMN engine, expert decisions	After initial assessment; upon route revision
Execution Block	Involved participants, resources, list of implemented measures, implementation timelines	Operational systems, BPMN tasks, case records, contractor reports	During execution
Re-verification Block	Results of post-remediation control, comparison of "before/after" states	Repeated monitoring, GIS, CV pipeline, laboratory data	After completion of measures
Current Status of Territory Block	Current state of the site, residual risk, usage restrictions	BPMN closure logic, monitoring layer, expert assessment	Upon case closure; upon repeated update
Archival-Documentary Block	Versioning, chronology of updates, case opening/closure/reopening	BPMN engine, event log, metadata layer	At all stages of the life cycle
Historical-Analytical Block	Versioning, history of changes, case status, analytical notes and conclusions	Event log, metadata layer, process mining outputs, expert analysis	At all stages; especially upon case closure

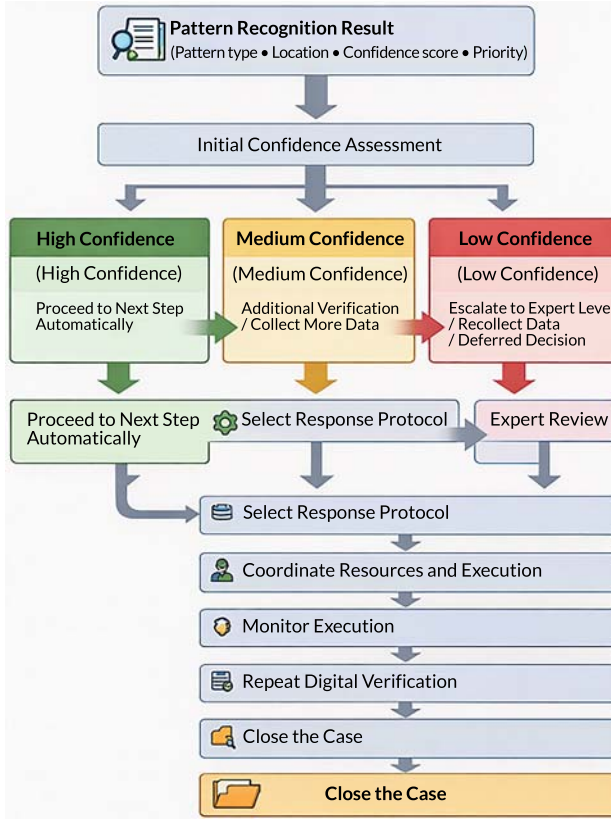


Fig. 7.2 Routing scheme of pattern recognition results by confidence level

This scheme reflects the key principle of adaptive support of Pattern Recognition results, which consists in the fact that the subsequent routing of the case is determined not only by the type of identified pattern, but also by the level of confidence of the digital output. In this sense, the confidence score acts as a central process trigger that determines the nature of subsequent managerial actions.

Thus, at a high level of confidence, the recognition result may serve as a basis for an automated transition to the selection of a response protocol, coordination of allocated resources for the elimination of consequences of the emergency event, and the initiation of remediation/recovery actions. A medium level of confidence reflects an intermediate zone of uncertainty, within which procedures of additional verification

are initiated – re-analysis of data, involvement of alternative data sources, as well as additional data collection and comparison with the geocontext. After that, the case is transferred either to the automated or to the expert branch. A low level of confidence excludes the automatic initiation of actions and requires urgent response: expert escalation, repeated data collection and/or additional diagnostics, that is, the transition of the process into a more controlled mode. Regardless of the initial routing branch, the process must remain closed-loop, and the result – strategically oriented toward priorities. After the selection of the response protocol, coordination of resources, and execution of the complex of remediation measures, monitoring must be carried out, followed by repeated digital verification, and only then – case closure. This makes it possible to shift management from the mode of reaction to a signal to the mode of confirmation of the achieved effect, which is especially important for remediation and post-crisis recovery.

Thus, the scheme not only illustrates branching by confidence levels, but also fixes the central idea of the study: the recognition result must be supported in a process-driven and adaptive manner up to the moment of confirmed case completion.

7.4 Synthesis of open benchmark foundations for verification of the proposed model

It should be noted that the panoramic scientific exploration conducted by the authors of the monograph revealed that no direct separate benchmark specifically for BPMN orchestration in remediation was identified. However, a strategy for proving the authors' concept through a composite benchmark consisting of three open contours is entirely realistic: a benchmark for damage recognition, open crisis geodata, and benchmark/event logs for process mining.

Thus, the three-layer composite benchmark consists of:

– a benchmark for damage/contamination recognition;

– an open contour for crisis/recovery geomonitoring;

– a benchmark/event log for process orchestration, adaptation, and execution analysis.

Direct Layer for Pattern Recognition. One of the most relevant benchmarks is xBD/xView2. xBD is positioned as a large open dataset for change detection and building damage assessment in humanitarian response and disaster recovery. Thus, the article on xBD [11] directly emphasizes the connection with damage assessment, logistics/resource planning, and post-crisis use. The xView competition series by DIU (Defense Innovation Unit) was generally created as a series of competitions

that benchmark computer vision algorithms for real humanitarian and disaster tasks, which correlates with the input layer "pattern recognition results" [32].

The second comprehensive source is the description of SpaceNet 8 [33–35]. It was collected as a benchmark for identifying flooded roads and buildings, and combines building detection, road extraction, and flood detection; the challenge provides a unified scoring logic for buildings and roads. For the authors' hypothesis, this is especially important because it makes it possible to move from object recognition/damage identification to the need to orchestrate recovery operations and logistics. For the crisis and recovery layer, there is a powerful open operational contour – Copernicus Emergency Management Service. Copernicus EMS is a service that provides mapping support for natural hazards, man-made emergencies, and humanitarian crises, and among the types of activations there is not only emergency response, but also recovery. It has damage assessment logic, methods of interpretation based on remote sensing, and open vector geodata. This is especially important for confirming the authors' concept, because it makes it possible to connect visual diagnosis with subsequent geospatial actions and managerial routes [16, 36–38].

Thus, if xBD/SpaceNet 8 provides the input diagnostic benchmark, then Copernicus EMS provides the operational-geospatial reference layer on which it is possible to model how BPMN transforms a recognition result into processes of verification, routing, notification, resource allocation, and subsequent closure. As a methodological benchmark for process evaluation, the official resources of the Process Mining Task Force may be used, as they publish event logs and BPI Challenge datasets, while the BPI Challenges themselves constitute a recognized benchmark class for analyzing real event logs. There are also catalogs of event data / logs for process mining. This means that the layer of "how to measure process behavior, deviations, delays, rework, bottlenecks, escalations" can rely on an already existing benchmark culture. That is, they confirm that the process layer can be evaluated according to the rules of process mining.

A separate block consists of a number of scientific sources confirming that in crisis situations and emergency response, namely in unpredictable and weakly structured crisis scenarios, procedural BPMN alone is often insufficient and the combination of BPMN + CMMN + DMN is useful. This is especially important for confirming the authors' hypothesis, because the proposed approach essentially requires:

– a procedural framework;

– flexibility by case;

– a decision layer [8, 39–41].

Thus, BPMN is the core of orchestration, while adaptive support is implemented through the composition of BPMN + rules/DMN + HITL + external services.

7.5 Possibilities of operationalization and verification of the architectural model

The architectural model proposed in the monograph is not an exclusively theoretical construction, since its key elements can be operationalized on the basis of open datasets, process modeling standards, and modern BPM (Business Process Management) platforms. In methodological terms, this is not a full-scale empirical approbation of an implemented system, but a proof of evaluability, that is, evidence that this concept can be formalized, implemented in an executable contour, and verified on open benchmark foundations. As a basic scenario for applying the model, the following chain may be considered: damage assessment based on xBD/xView2 data → formation of a structured event → launch of a BPMN contour with confidence-based routing → passage through one of three branches → selection of a response protocol to contamination and/or damage → monitoring of the remediation scenario execution → repeated digital verification → case closure.

At the first stage, the CV pipeline processes pre- and post-event images and forms a primary diagnostic output, including the type, class, and scale of contamination/damage, spatial reference of the object, and a confidence-like indicator. However, this output is not yet a managerial decision, since it must be transformed into a structured BPMN event that contains the case identifier, coordinates of the object/territory, pattern type, confidence level, list of data sources, and key priorities. It is precisely such an event that becomes the input for the executable remediation process contour.

Next, a BPMN process is initiated, within which an initial assessment of the criticality and confidence level of the verification result is carried out. It should be noted that at this stage a decision-making layer based on DMN may be used, since this standard is specifically intended for the formalization of repeatable decisions with the possibility of their automated execution. Accordingly, rules of the type "if the damage class of the object/territory is severe and the confidence level ≥ the established threshold, then initiate the automatic route of emergency consequence elimination" or "if the confidence level is within an intermediate range, then route this case for additional verification" may be applied. In the proposed model, as described earlier, it is advisable to provide for three routing branches, namely: a high level of confidence, which assumes the possibility of automatic transition to the selection of an emergency consequence elimination protocol/remediation scenario; a medium level of confidence, which assumes additional verification through repeated data analysis, aggregation of alternative information sources, as well as manual verification of attributes or refinement of the geocontext; a low level of confidence, which provides

for the transition of the case into the mode of expert review, repeated data collection if necessary, or deferred decision until confirmation is obtained.

Further, after the route selection, the system proceeds to the assignment of a response protocol/remediation scenario, where the previously described damage assessment service of the Copernicus Emergency Management Service may be used as a reference operational layer. Then, BPMN ensures full coordination of all participants of the remediation process within the implementation of the specifically selected scenario by means of service tasks for the purpose of intersystem information exchange and coordination of the execution of subprocesses and tasks for the elimination of the consequences of the emergency situation. After the initiation of the remediation scenario implementation, the architecture must provide for monitoring of the status and repeated digital verification of the result through the acquisition of new data, repeated recognition, or updating of the geospatial layer and comparison of the "before/after" state. Consequently, this step transforms the process into a closed-loop cycle of comprehensive management of the remediation process.

For the purpose of verifying this architectural model, it is advisable to use not domain-specific remediation event logs, which are currently practically absent in open access, but the methodological logic of process mining benchmark culture. In this case, it is not a specifically selected remediation scenario that will be evaluated, but the parameters of process execution, namely: lead time, rework rate, escalation frequency, expert intervention rate, false auto-routing rate, closure time, and post-verification success rate. Such a set of metrics makes it possible to assess how effectively the model manages the level of uncertainty of AI outputs and how rationally it distributes cases between automatic and expert branches. From a practical point of view, the most realistic implementation stack includes: CV pipeline + BPMN/DMN engine + persistence/event layer + monitoring/analytics layer.

Thus, as a CV pipeline, a damage assessment model trained or fine-tuned on xBD/xView2 may be used. As the process core, Camunda 8 or Flowable may be applied, since both platforms support BPMN and DMN, while Flowable is additionally oriented toward the integration of BPMN/CMMN/DMN within a unified set of engines. Camunda is particularly convenient in scenarios where confidence thresholds, criticality levels, and response protocols are defined not within the process code, but in the decision layer through DMN and FEEL (Friendly Enough Expression Language). Flowable, in turn, is of interest for further expansion of the model toward case-oriented logic and less structured scenarios.

In other words, adaptive support here does not require a special set of elements, since it emerges from the meaningful combination of events, gateways,

tasks, decision management, as well as exception-handling mechanisms. The recommended structure of the technical stack is presented below in **Table 7.3**.

Table 7.3 Recommended technical stack for implementation

Layer	Recommended Components	Purpose
Visual diagnosis layer	xBD/xView2-trained CV model; satellite imagery preprocessing	Building damage assessment, class prediction, confidence output
Event formalization layer	Python/Java service, API gateway, JSON event schema	Transformation of AI output into a structured BPMN
Process orchestration layer	Camunda 8 or Flowable BPMN engine	Execution of the BPMN routing contour
Decision layer	DMN engine, FEEL expressions, decision tables	Threshold rules, selection of response route and protocol
Case / expert layer	User tasks, task lists, notification service	HITL, expert verification, escalation
Operational geospatial layer	Copernicus EMS products, GIS services, map layers	Spatial context, recovery/damage reference
Persistence and integration layer	Event store, relational DB, REST / message broker	Storage of cases, statuses, results, and intersystem exchange
Monitoring and analytics layer	Process monitoring dashboard, event log export, process mining tools	Lead time, rework, escalation, closure and post-verification metrics

It should be emphasized that the above-presented stack describes an implementable, but not the only possible platform configuration. Its significance lies primarily in confirming the consistency of the authors' architecture, since it translates the concept into the plane of operationalization: it shows what data are required, which elements must be formalized, where the decision logic is located, and how the evaluation of process effectiveness may be organized.

7.6 Conclusion

The conducted study makes it possible to conclude that BPMN in the context of remediation and post-crisis recovery should be considered much more broadly – as a process framework for adaptive support of Pattern Recognition results and digital diagnostics. The authors substantiate that the key problem of modern digital management contours of emergency situations of natural, technogenic, and military nature lies in the gap between the obtained diagnostic conclusion and its transformation

into a sequence of verifiable, coordinated, and controllable managerial actions for consequence elimination and subsequent remediation of the object/territory.

The proposed concept forms an integral architectural model in which the recognition result is transformed into a structured event, passes through a BPMN contour of confidence-based routing, is complemented by decision logic and expert verification, and is brought to the stage of execution of the remediation scenario, repeated digital verification, and case closure. Thus, it is substantiated that adaptive support of AI outputs may be considered as an independent process mode of managing the level of uncertainty and reducing the risk of erroneous automation. An important outcome of the study is the conclusion regarding the fundamental operationalizability of the proposed model based on open benchmark datasets of visual diagnostics, geospatial recovery products, and process mining metrics. At the same time, the analysis of literature and practice has confirmed the presence of a study gap: despite the high maturity of individual directions – BPMN, disaster damage assessment, recovery mapping, and process evaluation – their integration into a unified contour of adaptive support of digital diagnostics results for remediation tasks remains insufficiently developed.

The limitations of the concept consist in its predominantly architectural and conceptual nature, the absence of full-scale empirical approbation, and the fact that existing benchmark foundations make it possible to verify only individual contours of the model. The prospects for further study are associated with scenario-based testing of the model on real crisis datasets, the construction of domain-specific event logs for process mining, testing of confidence-based routing rules, as well as the creation of a pilot BPMN/DMN prototype integrated with geospatial and computer vision services.

Conflict of interest

The authors declare that there is no conflict of interest in relation to this paper, as well as the published research results, including the financial aspects of conducting the research, obtaining and using its results, as well as any non-financial personal relationships.

Financing

The study was performed without financial support.

Data availability

The data underlying the conclusions of this study, including the used open benchmark datasets (xBD/xView2, SpaceNet 8, Copernicus EMS) and methodological materials, may be provided by the authors upon reasonable request.

Use of Artificial Intelligence statement

The authors employed artificial intelligence technologies solely for auxiliary purposes that do not affect the scientific novelty or reliability of the results. In particular, ChatGPT (OpenAI) was used for verifying the academic English translation and generating illustrative materials. All conceptual, methodological, and architectural contributions, as well as the study's conclusions, were developed exclusively by the authors.

Acknowledgments

The authors express their gratitude to the reviewers of this study and the publisher for providing the opportunity to publish the results of their scientific activity.

Authors' contributions

Bohdan Cherniavskyi: Conceptualization of the study, Development of the theoretical and methodological framework, Formation of the architectural model of adaptive support, Formal analysis, Preparation of the initial manuscript draft, Scientific editing, Integration of the research results.

Oksana Drozd: Methodology, Analysis of literary sources, Participation in the development of the research logic, Interpretation of theoretical provisions, Preparation of individual manuscript sections, Text editing.

Anatolii Nadtochyi: Methodology, Formal analysis, Participation in the development of the logic of model operationalization, Analysis of benchmark foundations, Participation in the formation of the concept of the Digital Passport of the territory, Validation of results, Editing of the final version of the text.

Viktor Nadtochii: Methodology, Development of the technical logic of operationalization and verification of the architectural model, Formation of the technical

implementation stack, Visualization, Preparation of diagrams, Tables, and Architectural solutions, Participation in the writing and editing of the manuscript.

Maksym Matviienko: Analysis of data and sources, Research on process mining and process orchestration in the applied context, Systematization of process metrics and open benchmark ecosystems, Participation in the development of model verification scenarios, Preparation of visualizations as well as analytical materials for the manuscript.

Ivan Kalinichenko: Research on the applied aspects of the context of remediation and post-crisis recovery, Analysis of geospatial data and recovery mapping, Participation in the development of practice-oriented provisions of the study, Validation of the substantive structure of the work, Development of visual content, Editorial refinement, and Approval of the final version of the manuscript.

References

1. Global Assessment Report on Disaster Risk Reduction 2025: Resilience Pays: Financing and Investing for Our Future (2025). United Nations Office for Disaster Risk Reduction. Stylus Publishing, LLC. Available at: https://www.undrr.org/gar/gar2025
2. Earth Observations for Disaster Risk Assessment & Resilience (2019). NASA Applied Remote Sensing Training Program (ARSET). Available at: https://www.earthdata.nasa.gov/learn/trainings/earth-observations-disaster-risk-assessment-resilience
3. Scheip, C. M., Wegmann, K. W. (2021). HazMapper: a global open-source natural hazard mapping application in Google Earth Engine. Natural Hazards and Earth System Sciences, 21(5), 1495–1511. https://doi.org/10.5194/nhess-21-1495-2021
4. Humanitarian Applications Using NASA Earth Observations (2022). NASA Applied Remote Sensing Training Program (ARSET). Available at: https://www.earthdata.nasa.gov/learn/trainings/humanitarian-applications-using-nasa-earth-observations
5. Chawanji, S., Fleischer, K., Ullrich, J. (2022). Copernicus Emergency Management Service (CEMS) – Risk and Recovery Mapping. Abstracts of the ICA, 5, 1–2. https://doi.org/10.5194/ica-abs-5-118-2022
6. von Rosing, M., White, S., Cummins, F., de Man, H. (2015). Business Process Model and Notation – BPMN. The Complete Business Process Handbook. Elsevier, 433–457. https://doi.org/10.1016/b978-0-12-799959-3.00021-5

7. Cherniavskyi, B., Cherniavska, T. (2024). Application of BPMN for management of remediation processes in war-affected territories of Ukraine. Kontrol i upravlinnia v skladnykh systemakh (KUSS-2024). Available at: https://conferences.vntu.edu.ua/index.php/mccs/mccs2024/paper/viewFile/22980/19039
8. Niemz, S., Gehrke, S., Ruhland, J. (2021). On process organization in crisis situations with BPMN, CMMN and DMN. 37th ibima conference. Available at: https://www.researchgate.net/profile/Sandra-Niemz/publication/351346718_On_Process_Organization_in_Crisis_Situations_with_BPMN_CMMN_and_DMN/links/60928530a6fdccaebd096d11/On-Process-Organization-in-Crisis-Situations-with-BPMN-CMMN-and-DMN.pdf
9. Cossentino, M., Guastella, D. A., Lopes, S., Sabatucci, L., Tripiciano, M. (2022). From Textual Emergency Procedures to Executable Plans. Proceedings of the 19th ISCRAM Conference. Tarbes. 200–212. Available at: https://idl.iscram.org/files/massimocossentino/2022/2410_MassimoCossentino_etal2022.pdf
10. Routis, I., Bardaki, C., Nikolaidou, M., Dede, G., Anagnostopoulos, D. (2022). Exploring CMMN applicability to knowledge-intensive process modeling: An empirical evaluation by modelers. Knowledge and Process Management, 30 (1), 33–54. https://doi.org/10.1002/kpm.1738
11. Gupta, R., Hosfelt, R., Sajeev, S., Patel, N., Goodman, B., Doshi, J. et al. (2019). xbd: A dataset for assessing building damage from satellite imagery. arXiv preprint arXiv:1911.09296. https://doi.org/10.48550/arXiv.1911.09296
12. Hafner, S., Gerard, S., Sullivan, J., Ban, Y. (2025). DisasterAdaptiveNet: A robust network for multi-hazard building damage detection from very-high-resolution satellite imagery. International Journal of Applied Earth Observation and Geoinformation, 143, 104756. https://doi.org/10.1016/j.jag.2025.104756
13. Su, J., Bai, Y., Wang, X., Lu, D., Zhao, B., Yang, H. et al. (2020). Technical Solution Discussion for Key Challenges of Operational Convolutional Neural Network-Based Building-Damage Assessment from Satellite Imagery: Perspective from Benchmark xBD Dataset. Remote Sensing, 12 (22), 3808. https://doi.org/10.3390/rs12223808
14. Gholami, S., Robinson, C., Ortiz, A., Yang, S., Margutti, J., Birge, C. et al. (2022). On the Deployment of Post-Disaster Building Damage Assessment Tools using Satellite Imagery: A Deep Learning Approach. 2022 IEEE International Conference on Data Mining Workshops (ICDMW). IEEE, 1029–1036. https://doi.org/10.1109/icdmw58026.2022.00134
15. Mitraka, Z., Siachalou, S., Doxani, G., Patias, P. (2020). Decision Support on Monitoring and Disaster Management in Agriculture with Copernicus Sentinel Applications. Sustainability, 12 (3), 1233. https://doi.org/10.3390/su12031233

16. Copernicus EMS On Demand Mapping. Available at: https://mapping.emergency.copernicus.eu
17. Risk and Recovery for Recovery Portfolio (2025). Copernicus EMS On Demand Mapping. Available at: https://mapping.emergency.copernicus.eu/about/recovery-mapping-portfolio/
18. Copernicus Emergency Management Service (EMS) (2012). United Nations Office for Outer Space Affairs. UN-SPIDER Knowledge Portal. Available at: https://www.un-spider.org/copernicus-emergency-management-service-ems
19. Gimenez, R., Lassalle, G., Elger, A., Dubucq, D., Credoz, A., Fabre, S. (2022). Mapping Plant Species in a Former Industrial Site Using Airborne Hyperspectral and Time Series of Sentinel-2 Data Sets. Remote Sensing, 14 (15), 3633. https://doi.org/10.3390/rs14153633
20. Agarwal, V., Kumar, M., Panday, D. P., Zang, J., Munoz-Arriola, F. (2024). Unlocking the potential of remote sensing for arsenic contamination detection and management: Challenges and perspectives. Current Opinion in Environmental Science & Health, 42, 100578. https://doi.org/10.1016/j.coesh.2024.100578
21. Hussain, A., Bangroo, S. A., Muslim, M.; Mushtaq, F., Farooq, M., Mukherjee, A. B., Ghosh Nee Lala, M. (Eds.) (2023). Assessment of Soil Contamination Using Remote Sensing and Spatial Techniques. Geospatial Analytics for Environmental Pollution Modeling. Cham: Springer, 249–266. https://doi.org/10.1007/978-3-031-45300-7_10
22. Dean, J. R., Ahmed, S., Cheung, W., Salaudeen, I., Reynolds, M., Bowerbank, S. L. et al. (2024). Use of remote sensing to assess vegetative stress as a proxy for soil contamination. Environmental Science: Processes & Impacts, 26 (1), 161–176. https://doi.org/10.1039/d3em00480e
23. Organisation. 4TU.ResearchData. Available at: https://community.data.4tu.nl/our-organisation/
24. Weytjens, H. (2021). predictive-process-monitoring-benchmarks. GitHub. Available at: https://github.com/hansweytjens/predictive-process-monitoring-benchmarks
25. Natarajan, S., Mathur, S., Sidheekh, S., Stammer, W., Kersting, K. (2025). Human-in-the-loop or AI-in-the-loop? Automate or Collaborate? Proceedings of the AAAI Conference on Artificial Intelligence, 39 (27), 28594–28600. https://doi.org/10.1609/aaai.v39i27.35083
26. Lazaros, K., Vrahatis, A. G., Kotsiantis, S. (2026). Human-in-the-Loop Artificial Intelligence: A Systematic Review of Concepts, Methods, and Applications. Entropy, 28 (4), 377. https://doi.org/10.3390/e28040377

27. Kovalerchuk, B. (2024). T02: Explainable AI for High-stakes tasks with Human-in-the-loop. HCI International 2024. Washington. Available at: https://2024.hci.international/T02.html
28. Hattab, G. (2026). Human Agency, Causality, and the Human Computer Interface in High-Stakes Artificial Intelligence. arXiv preprint arXiv:2604.12793. Available at: https://arxiv.org/html/2604.12793v1
29. Business Process Model & Notation™ (BPMN™). Object Management Group. Available at: https://www.omg.org/bpmn/
30. Business Process Model and Notation (BPMN). Version 2.0.2 (2013). Object Management Group. Available at: https://www.omg.org/spec/BPMN/2.0.2/PDF
31. Delgado, A., Calegari, D. (2019). Towards integrating BPMN 2.0 with CMMN and DMN standards for flexible business process modeling. CIbSE. La Habana, 697–704. Available at: https://www.researchgate.net/publication/345347184_Towards_integrating_BPMN_20_with_CMMN_and_DMN_standards_for_flexible_business_process_modeling
32. xView Challenge Series. Defense Innovation Unit. Available at: https://www.diu.mil/ai-xview-challenge
33. Hänsch, R., Arndt, J., Lunga, D., Gibb, M., Pedelose, T., Boedihardjo, A. et al. (2022). SpaceNet 8 – The Detection of Flooded Roads and Buildings. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 1472–1480. Available at: https://openaccess.thecvf.com/content/CVPR2022W/EarthVision/papers/Hansch_SpaceNet_8_-_The_Detection_of_Flooded_Roads_and_Buildings_CVPRW_2022_paper.pdf
34. Arndt, J., Lunga, D. (2022). SpaceNet 8 – The Detection of Flooded Roads and Buildings. OSTI.GOV. Available at: https://www.osti.gov/biblio/1996730
35. SpaceNet8 (2023). SpaceNet Challenge. GitHub. Available at: https://github.com/SpaceNetChallenge/SpaceNet8
36. Detection methods and Damage Assessment (2025). Copernicus EMS On Demand Mapping. Available at: https://mapping.emergency.copernicus.eu/about/rapid-mapping-manual/detection-methods-damage-assessment/
37. Vector Data (2024). Copernicus EMS On Demand Mapping. Available at: https://mapping.emergency.copernicus.eu/about/risk-and-recovery-manual/downloadable-resources/geodata/data-formats/vector-data/
38. Datasets: Copernicus EMS. JRC Data Catalogue. Available at: https://data.jrc.ec.europa.eu/dataset?keyword=Copernicus+EMS
39. Ruiz Herrera, M. P., Sánchez Díaz, J.; Franco, Z., Gonzáles, J. J., Canós, J. H. (Eds.) (2019). Improving Emergency Response through Business Process,

Case Management, and Decision Models. Proceedings of the 16th ISCRAM Conference. València: ISCRAM, 116–125. Available at: https://idl.iscram.org/files/marcopoloruizherrera/2019/1944_MarcoPoloRuizHerrera%2BJuanSanchezDiaz2019.pdf

40. Shahrah, A. Y., Al-Mashari, M. A. (2017). Modelling emergency response process using case management model and notation. IET Software, 11 (6), 301–308. Portico. https://doi.org/10.1049/iet-sen.2016.0209
41. Bule, M., Polančič, G. (2025). Analysis and Synthesis of Theoretical and Practical Implications of Case Management Model and Notation. Information, 16 (4), 310. https://doi.org/10.3390/info16040310