НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ
НАВЧАЛЬНО-НАУКОВИЙ КОМПЛЕКС
«ІНСТИТУТ ПРИКЛАДНОГО СИСТЕМНОГО АНАЛІЗУ»
НАЦІОНАЛЬНОГО ТЕХНІЧНОГО УНІВЕРСИТЕТУ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО»

СИСТЕМНІ ДОСЛІДЖЕННЯ ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

МІЖНАРОДНИЙ НАУКОВО–ТЕХНІЧНИЙ ЖУРНАЛ

№ 2 2025

ЗАСНОВАНО У ЛИПНІ 2001 р.

АДРЕСА РЕДАКЦІЇ:

03056, м. Київ,
просп. Перемоги, 37, корп. 35,
ННК «ІПСА» КПІ ім. Ігоря Сікорського
Тел.: **204-81-44**; факс: **204-81-44**
E-mail: journal.iasa@gmail.com
http://journal.iasa.kpi.ua

У номері:

NATIONAL ACADEMY OF SCIENCES OF UKRAINE
EDUCATIONAL AND SCIENTIFIC COMPLEX
«INSTITUTE FOR APPLIED SYSTEM ANALYSIS»
OF THE NATIONAL TECHNICAL UNIVERSITY OF UKRAINE
«IGOR SIKORSKY KYIV POLYTECHNIC INSTITUTE»

SYSTEM RESEARCH AND INFORMATION TECHNOLOGIES

INTERNATIONAL SCIENTIFIC AND TECHNICAL JOURNAL

№ 2 2025

IT IS FOUNDED IN JULY 2001

THE EDITION ADDRESS:

03056, Kyiv,
av. Peremogy, 37, building 35,
Institute for Applied System Analysis
at the Igor Sikorsky Kyiv Polytechnic Institute
Phone: **204-81-44**; Fax: **204-81-44**
E-mail: journal.iasa@gmail.com
http://journal.iasa.kpi.ua

In the issue:

- **Progressive information technologies, high-efficiency computer systems**
- **Theoretical and applied problems of intelligent systems for decision making support**
- **Mathematical methods, models, problems and technologies for complex systems research**
- **Scientific and methodological problems in education**

Шановні читачі!

Навчально-науковий комплекс «Інститут прикладного системного аналізу» Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського» видає міжнародний науково-технічний журнал

«СИСТЕМНІ ДОСЛІДЖЕННЯ ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ».

Журнал публікує праці теоретичного та прикладного характеру в широкому спектрі проблем, що стосуються системних досліджень та інформаційних технологій.

Провідні тематичні розділи журналу:

Теоретичні та прикладні проблеми і методи системного аналізу; теоретичні та прикладні проблеми інформатики; автоматизовані системи управління; прогресивні інформаційні технології, високопродуктивні комп'ютерні системи; проблеми прийняття рішень і управління в економічних, технічних, екологічних і соціальних системах; теоретичні та прикладні проблеми інтелектуальних систем підтримання прийняття рішень; проблемно і функціонально орієнтовані комп'ютерні системи та мережі; методи оптимізації, оптимальне управління і теорія ігор; математичні методи, моделі, проблеми і технології дослідження складних систем; методи аналізу та управління системами в умовах ризику і невизначеності; евристичні методи та алгоритми в системному аналізі та управлінні; нові методи в системному аналізі, інформатиці та теорії прийняття рішень; науково-методичні проблеми в освіті.

Головний редактор журналу — науковий керівник ННК «ІПСА» КПІ ім. Ігоря Сікорського, академік НАН України Михайло Захарович Згуровський.

Журнал «Системні дослідження та інформаційні технології» включено до переліку наукових фахових видань України (категорія «А»).

Журнал «Системні дослідження та інформаційні технології» входить до таких наукометричних баз даних: Scopus, EBSCO, Google Scholar, DOAJ, Index Copernicus, реферативна база даних «Україніка наукова», український реферативний журнал «Джерело», наукова періодика України.

Статті публікуються українською та англійською мовами.

Журнал рекомендовано передплатити. **Наш індекс 23918.** Якщо ви не встигли передплатити журнал, його можна придбати безпосередньо в редакції за адресою: 03056, м. Київ, просп. Перемоги, 37, корп. 35.

Завідувачка редакції **С.М. Шевченко**
Редакторка **Р.М. Шульженко**
Молодша редакторка **Л.О. Тарин**
Комп'ютерна верстка, дизайн **А.А. Патіохи**

Рішення Національної ради України з питань телебачення і радіомовлення
№1794 від 21.12.2023. Ідентифікатор медіа R30-02404

Підписано до друку 28.06.2025. Формат 70×108 1/16. Папір офс. Гарнітура Times.
Спосіб друку – цифровий. Ум. друк. арк. 14,411. Обл.-вид. арк. 28,56. Наклад 153 пр. Зам. № 11/04

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Свідоцтво про державну реєстрацію: ДК № 5354 від 25.05.2017 р.
просп. Перемоги, 37, м. Київ, 03056.
ФОП Пилипенко Н.М., вул. Мічуріна, б. 2/7, м. Київ, 01014. тел. (044) 361 78 68.
Виписка з Єдиного державного реєстру № 2 070 000 0000 0214697 від 17.05.2019 р.

Dear Readers!

Educational and Scientific Complex «Institute for Applied System Analysis» of the National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute» is published of the international scientific and technical journal

«SYSTEM RESEARCH AND INFORMATION TECHNOLOGIES».

The Journal is printing works of a theoretical and applied character on a wide spectrum of problems, connected with system researches and information technologies.

The main thematic sections of the Journal are the following:

Theoretical and applied problems and methods of system analysis; theoretical and applied problems of computer science; automated control systems; progressive information technologies, high-efficiency computer systems; decision making and control in economic, technical, ecological and social systems; theoretical and applied problems of intellectual systems for decision making support; problem- and function-oriented computer systems and networks; methods of optimization, optimum control and theory of games; mathematical methods, models, problems and technologies for complex systems research; methods of system analysis and control in conditions of risk and uncertainty; heuristic methods and algorithms in system analysis and control; new methods in system analysis, computer science and theory of decision making; scientific and methodical problems in education.

The editor-in-chief of the Journal is scientific director of the Institute for Applied System Analysis at the Igor Sikorsky Kyiv Polytechnic Institute, academician of the NASU Michael Zaharovich Zgurovsky.

The articles to be published in the Journal in Ukrainian and English languages are accepted. Information printed in the Journal is included in the Catalogue of periodicals of Ukraine.

СИСТЕМНІ ДОСЛІДЖЕННЯ ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

2 • 2025

ЗМІСТ

SYSTEM RESEARCH AND INFORMATION TECHNOLOGIES

2 • 2025

CONTENT

UDC 004.942: 303.732.4
DOI: 10.20535/SRIT.2308-8893.2025.2.01

FROM CAD AND BIM TO DIGITAL TWINS

A.I. PETRENKO

Abstract. A Digital Twin is a virtual model of a physical object or system that uses real-time data to simulate the behavior of its real counterpart. It can be a product, machine, building, or even an entire city. *Digital modeling* is a fundamental technology for creating Digital Twins, as it provides a methodology for representing physical objects in the virtual world. *CAD (Computer-Aided Design)* can be used to create the initial model of a Digital Twin. *BIM (Building Information Modeling)* is a specialized form of CAD that focuses on building projects incorporating more information than just geometry but also data on time, costs, operations, and maintenance. This paper examines how these technologies are increasingly integrated with each other, utilizing *mathematical modeling* that, through virtual computational experiments, provides an understanding of the complex functioning of objects and informed decision-making in various fields. The integration of Digital Twins and CAD is transforming the ways products are designed, modeled, and optimized in the industry. BIM models can serve as the basis for creating Digital Twins of buildings, which are then used to optimize energy consumption, maintenance, and repair. With the growth of the Internet of Things (IoT), Digital Twins are receiving more and more real data, making them even more accurate and useful for forecasting and optimization. The use of artificial intelligence to analyze data collected by digital twins allows predicting breakdowns, optimizing processes, and even automating the control of systems.

Keywords: mathematical modeling, digital modeling, CAD and BIM, Digital Twins (DT), Internet of Things (IoT), AI application in DT.

MATHEMATICAL MODELING (SIMULATION)

Among the information technologies that are used in almost all fields of engineering and science, *digital modeling* occupies a special place. This is a general term that encompasses the creation of digital representations of physical objects or systems. It can include many things: from modeling their functioning to creating their geometric images. Therefore, *mathematical modeling* is separately distinguished, when under the mathematical model of a physical system, object, or process is usually understood a set of mathematical relations (formulas, equations, logical expressions) that determine the characteristics of the state and properties of the system, object, and process and their functioning depending on the parameters of their components, initial conditions, input disturbances, and time. In general, a mathematical model describes the functional relationship between the output dependent variables, through which the functioning of the system is reflected, independent (such as time), and changing variables (such as component parameters, geometric dimensions, etc.), as well as input disturbances applied to the system.

Mathematical models are determined by the subject area of design:

- physical and mathematical description of the laws and conditions of the object's functioning;
- the environment of functioning and means of interaction of the object with this environment;
- the composition of the object, the element base, the means of organizing the structure of the object;
- parameters that change or are adjusted.

The models distinguish between three types of data: data on the elements of the modeling object; data on the properties of the object and elements; data on the relationship between elements and the object. The abstraction of the object is carried out both by *the depth of structuring* (hierarchical system, system of elements, indivisible object), and by *the degree of abstraction of elements* and object properties (structural, logical, and quantitative levels).

At the structural level, the structure of the object is modeled at the lowest level of its structuring, when the mathematical model is presented in the form of a set, the properties and parameters of the elements of which are described through functional connections, order relations, adjacency, combination. In this case, the apparatus of set theory and graphs, queuing theory, etc. is used.

At the logical level of modeling, each set, Boolean matrix of binary relations, or structural graph corresponds to sets of logical relations and variables that reflect cause-and-effect relationships. In design, the apparatus of mathematical logic is used.

At the quantitative level, each element of the set, Boolean matrix, or logical variable is assigned an algebraic or other quantitative variable, and logical relations are transformed into quantitative relations: equations, inequalities, and so on. Modeling at the quantitative level reflects functional, energy, material, and spatial connections. These connections are usually described by spatio-temporal relations and are determined through ordinary differential equations or partial differential equations.

If the object consists of a system of elements, the connection between which is described by only one variable (time), then models with lumped parameters are used. The elements of the object are quantitatively described by component equations. Micro models and macro models are distinguished if the internal structure of the modeled objects or elements is not taken into account.

The main part of the calculations in mathematical modeling (in terms of volume and costs) is performed at the quantitative level of modeling, where all existing non-linear relations are taken into account. Modern modeling programs differ from the simple use of computers in calculations by the fact that they provide ***automatic formation of a digital mathematical model*** of an object based on information about its structure and the properties of its elements [1]. For example, for the common case of dynamic non-linear systems, the components of which can be electronic blocks (logical and analog), mechanical, hydraulic, pneumatic, electromagnetic components, digital mathematical models of the object in the domestic complex ALLTED (**ALL TE**chnology **D**esigner) are described by joint systems of algebra-differential equations or only differential equations [2]. At the same time, all stages of designing non-linear dynamic systems are supported: constructing a mathematical model of the object, analyzing direct current (DC), analyzing time dynamics (TR) and frequency properties (AC), statistical analysis (Monte-Carlo), Fourier analysis (Four), worst-case analysis (WCD), sensitivity analysis (SA), optimization of parameters and characteristics (OPTIM), optimal assignment of tolerances (TOLAS), etc. This virtual laboratory is based on client-server technology and allows serving many clients located in different cities and countries.

It differs from existing foreign analogues (Pspice (Microsim Corp. USA), Saber (Analogy Corp. USA)) by:

- original algorithms of numerical procedures that allow solving "stiff" and insufficiently conditioned problems in DC, TR, SA modes;
- original algorithms for optimizing variable orders that take into account not three (as usual), but five terms in the equivalent Taylor series without calculating derivatives higher than the second order;
- powerful procedures for automatically calculating design parameters (time delays, rise and fall times of pulses, resonance frequency or bandwidth, power consumption, etc.) and functions of these parameters;
- the possibility of solving single- and multi-criteria optimization problems with parametric and functional constraints, and the optimized variables can be both primary parameters (voltages, currents, powers) and the above-mentioned design parameters;
- the possibility of using user-defined component models along with the use of powerful libraries of models and parameters;
- the procedure of optimal assignment of tolerances to the parameters of internal components, based on a given allowable deviation of the values of output parameters or variables, which allows organizing diagnostics of malfunctions in the structure of the object.

Such possibilities of modern modeling complexes allow in many cases to abandon the physical prototyping of designed products, replacing it with mathematical modeling (computational experiment), which is especially important when physical prototyping is difficult or practically impossible (for example, modeling a dam break, moving a rover on the surface of Mars, etc.). These innovative toolkits have allowed humanity to reach the current level of development, in particular, to create modern microchips (as an element base of Industry 4.0), powerful rocket technology, modernize the capabilities of mechanical engineering, energy, agriculture, industrial and civil construction, and make space flights and research.

COMPUTER-AIDED DESIGN (CAD)

CAD (Computer-Aided Design) systems appeared in the early 1960s and are a form of digital modeling, where the emphasis is on design and technical details. The term was first introduced in the late 1950s by *Dr. Patrick Hanratty*. He is often called the "father of CAD", and he was responsible for creating PRONTO, the software that started CAD. But AutoCAD, the first commercially available drafting software, was released in 1982. CAD is used to create 3D models and 2D models of objects. These models, which contain detailed information about the geometry, materials, and other characteristics of the object, can be of varying complexity and detail, from simple parts to complex mechanisms and buildings.

Although CAD and mathematical modeling are closely related, there are some key differences between them [1]:

Focus:

- *CAD:* the main focus of CAD is on the geometric representation of the object. It is used to create accurate 2D drawings and 3D models, with an emphasis on visualization, design, and documentation.
- *Mathematical modeling:* focuses on the mathematical description of the behavior of a system or process. It uses mathematical equations, formulas, and algorithms for analysis, forecasting, and optimization.

Tools:

- *CAD:* uses specialized software with tools for creating and editing geometric shapes, dimensions, annotations, etc. Examples: AutoCAD, SolidWorks, CATIA.
- *Mathematical modeling*: uses a variety of tools, including programming languages (Python, MATLAB), mathematical packages (Mathematica, Maple), specialized modeling software (Simulink, AnyLogic).

Results:

- *CAD*: the result is a graphical representation.
- *Mathematical modelling*: the result is a mathematical description of the system, forecasts, sensitivity analysis, and optimization solutions.

Applications:

- *CAD:* Widely used in engineering, architecture, and design for the design and development of products, buildings, and mechanisms.
- *Mathematical modeling*: Applied in various fields, including physics, economics, biology, and finance, for the analysis and prediction of complex systems and processes.

CAD and mathematical modeling can complement each other. For example, a CAD model can be used as a basis for mathematical modeling, providing geometric data and parameters. Mathematical modeling, in turn, can help in the analysis and optimization of a structure created in CAD. In other words: CAD is like drawing a detailed portrait of an object, and mathematical modeling is like writing an equation that describes how this object moves or functions.

CAD is an excellent tool for engineers not only at the stage of optimal design, it also allows preparing high-quality design and technological documentation for the manufacture of designed objects, as well as diagnosing and testing manufactured products. Initially developed for two-dimensional design, CAD has evolved into powerful software for three-dimensional modeling, providing accurate digital modeling for the design and testing of structures and infrastructure. High-quality work with CAD often requires powerful computer systems. Mastering CAD software may require significant training and practice. Premium CAD programs can be expensive, and licenses and updates add to the cost. Although there are free or cheaper alternatives, they may not offer the same wide functionality or compatibility. Therefore, CAD is usually not used for long-term maintenance and servicing of objects during their life cycle.

BUILDING INFORMATION MODELING (BIM)

In the 1970s, BIM (Building Information Modelling) technology appeared — a more advanced digital representation of the physical and functional characteristics of an object. BIM is the process of creating and managing digital information about a building throughout its life cycle. A BIM model contains detailed information about the geometry, materials, structures, systems, and equipment of the building. With the help of BIM technology, an information model is created that provides an accurate vision of the project as a whole. The development of ArchiCAD, one of the most popular BIM software products, began in 1982 under the leadership of *Gábor Bojár*.

One of the main advantages of BIM is the ability to visualize the object in three-dimensional space. This allows designers and customers to more accurately imagine the future object, as well as make changes and additions in real time. Visualization in BIM can be represented in various formats — from static images to interactive 3D models. The BIM model contains information about all components of the object — from structural elements to electrical equipment and plumbing. BIM allows integrating all data about the object into a single digital model, which becomes the basis for all stages of the object's life cycle — from design and construction to operation and reconstruction. The use of BIM technology in the design of houses includes the collection and complex processing of technological, architectural, structural, and economic information about the building, so that the building object and everything related to it are considered as a single whole.

A similar software product from Graphisoft, known in architectural design circles, is called BIMx and is used as an important addition to their main CAD program ArchiCAD.

Developing on the basis of the foundations laid by CAD, BIM complements operations on the object with geometric complex 3D models with a large amount of data. As with CAD, BIM has several key issues related to the efficiency and feasibility of its use for the daily management and operation of objects. Detailed, data-rich BIM models can be overwhelming, especially when only certain subsets of data are needed. BIM tools and practices can be complex and may require extensive training for professionals to become proficient. BIM software often requires high-performance computer equipment. Maintaining BIM models up-to-date, especially for long-term projects, can be resource-intensive.

The building information model is a virtual prototype of a building structure, so the use of BIM technology in the design of houses allows you to check and evaluate various solutions before the start of construction work. Today, there are many different solutions on the BIM software market that allow you to perform design in a three-dimensional format and use it at all stages of construction. In Ukraine, the most popular programs for BIM design are AutoCAD Architecture, Revit and Allplan Architecture [3–5].

The use of BIM technology in construction design makes every action transparent and provides complete control, and in automated mode, which guarantees high quality of design and construction work. Today, BIM is a standard tool for working in the construction industry, which allows optimizing the processes of design, construction and operation of facilities.

BIM also provides the ability to create and store digital documentation that contains all the necessary documents — from drawings and specifications to calculations and statistical data. This allows easy control of the design and construction process of the object, as well as improve communication between project participants. BIM technologies can also be used in the process of building a building. With their help, it is possible to monitor the execution of works in accordance with the project and specifications; optimize the construction process and manage the timing of the project; minimize the risk of errors and conflicts between project participants and reduce the amount of waste and material costs. BIM technologies can be used in the process of building operation. With their help, it is possible to create electronic maps of the building with information about each of its elements and systems; track changes and maintain the building in real time; optimize the use of building resources (energy, water, heat, etc.); manage the timing and costs of maintenance and repair of the building.

BIM can be used for project management, coordination, risk assessment and compliance with standards. However, there are many other opportunities to expand the functionality of BIM, including inventory management, quality management, construction planning, etc. Expanding the functionality of BIM will help make the construction process more efficient and transparent. The development of open standards and data exchange protocols can help eliminate this problem and ensure more flexible integration of BIM with other systems, such as the Internet of Things (IoT) technology.

The Internet of Things (IoT) is a technology that allows networked objects to exchange data with each other and with other systems. The integration of BIM with IoT allows you to receive real-time data on what materials are used, what conditions are on the construction site, etc. This will help to more effectively manage the construction process and prevent possible problems.

Virtual and augmented reality (VR and AR) can be integrated with BIM to create more accurate and realistic visualizations of the project. This will allow more effectively verify the project for compliance with requirements and interact with customers and other project participants. In addition, virtual and augmented reality technologies can help in training personnel and improving safety at the construction site. For example, virtual simulators can be used to train workers without risk to their lives and health. In general, the development of virtual and augmented reality technologies can lead to a large leap in the development of BIM technologies in the future, increasing their efficiency and accuracy, as well as simplifying their use.

When preserving architectural heritage, it is important to be able to carry out historical conservation and restoration of buildings in conditions when the original architectural and construction documentation is absent. Laser technology "Scan to BIM" [6] comes to the rescue, offering accurate 3D scanning of historical structures and artifacts, which allows creating detailed digital models that reliably reflect the intricacies of these objects. This technology allows restorers and conservation experts to carefully analyze the condition of heritage objects, identify areas that need attention, and plan restoration work with meticulous accuracy. In addition, digital preservation using BIM facilitates long-term management and documentation of historical objects, ensuring that their heritage is preserved for future generations who can appreciate and learn from it.

“Scan to BIM” is a revolutionary process that uses laser scanning technology to create a virtual copy of existing structures, from grand cathedrals to quaint cottages, when laser precision captures dimension, identifies materials, and even reveals structural elements hidden behind layers of paint and plaster. This data becomes the basis for a smarter and more efficient way to manage, reconstruct, and even redesign existing buildings. The “Scan to BIM” technology transforms not only repair work, but also how building management is carried out. Its digital model tracks performance, predicts problems before they become critical, and optimizes maintenance schedules. Imagine proactive repairs, extended equipment life, and reduced operating costs — all thanks to the foresight of BIM’s digital intelligence.

The leader in the implementation of “Scan to BIM” technology is Harmony AT, which seamlessly integrates scan data with BIM methodologies, offering comprehensive solutions tailored to the diverse needs of projects. The company’s services cover a wide range of applications, including reconstruction, modernization, clash detection and facility management. In addition, its commitment to quality assurance, transparent communication and customer satisfaction makes the company a reliable partner in the construction industry. Whether it is about reviving historical monuments or optimizing modern construction projects, Harmony AT’s scan-to-BIM services embody efficiency, accuracy and excellence.

DIGITAL TWINS

Digital Twins are a new generation solution based on the foundation laid by CAD and BIM. While CAD and BIM have made significant contributions to the design and manufacturing stages, Digital Twins aim to rethink how we interact with, maintain, and operate the digital environment we create. These are not just static images, but dynamic models that reflect real objects or systems in real time.

Digital Twins (DT) are virtual models of real objects or processes that reflect their characteristics and behavior in dynamics [7,8]. They appeared as a result of the development of information technologies, in particular mathematical modeling, CAD and BIM.

The history of Digital Twins’ development has several stages:

- *The origin of the concept (2002): Michael Grieves* first presented the concept of DT at a conference in the USA. He proposed creating virtual copies of physical objects to manage their life cycle.
- *First applications (2010s):* With the development of technology, the first real projects for the use of data centres appeared in the aerospace and manufacturing industries. They were used to model, optimise and predict the behaviour of complex systems.
- *Active distribution (2015-present):* Thanks to the development of the Internet of Things, artificial intelligence, and cloud technologies, data centres have become available for a wide range of applications. They are used in various industries, such as industry, energy, medicine, construction, and others.

A Digital Twin is a virtual representation of a physical object, system or process that covers its entire life cycle. It is constantly updated with real-time data and uses modelling, machine learning and artificial intelligence to optimise decision-making. This technology allows companies to analyse performance, predict failures, and improve overall efficiency.

The connection between CAD and Digital Twins is that CAD models serve as a starting point for creating Digital Twins [9]. A CAD model is static, that is, it reflects the object at a certain point in time. To create a Digital Twin that reflects the behavior of an object in dynamics, it is necessary to supplement the CAD model with information about the physical properties of materials, operating conditions, and other parameters. This information can be obtained from various sources, such as sensors, monitoring systems, and others.

Digital Twins created on the basis of CAD models can be used to solve various tasks, such as:

- *Predicting object behavior*: Digital Twins can be used to model the behavior of an object in different operating conditions. This allows predicting possible breakdowns and accidents, as well as optimizing equipment operating modes.
- *Process optimization:* Digital Twins can be used to optimize production processes, design new products, and other tasks.
- *Virtual training:* Digital Twins can be used to train personnel to work with complex equipment.

In general, CAD and Digital Twins are interconnected technologies that complement each other. CAD models are the basis for creating Digital Twins, and Digital Twins, in turn, expand the capabilities of CAD models, allowing modeling the behavior of objects in dynamics and solving various tasks.

A Digital Twin of a building, in turn, is an expanded version of a BIM model, which includes not only static information about the building, but also dynamic data about its operation, such as:

- Data from sensors (temperature, humidity, energy consumption).
- Information about the condition of systems (ventilation, heating, lighting).
- Data on the use of the building (number of people, traffic).

The relationship between BIM and Digital Twins:

- *BIM as a basis:* the BIM model is the basis for creating a Digital Twin of a building. It provides detailed information about the physical characteristics of the building, which is necessary to create a virtual model.
- *Dynamic data:* the Digital Twin complements the BIM model with dynamic data that is collected during the operation of the building. This allows creating a completer and more accurate picture of the building's condition.
- *Analysis and optimization*: the Digital Twin, created on the basis of the BIM model, can be used to analyze and optimize various aspects of building operation, such as energy efficiency, comfort, safety, and others.

Advantages of using Digital Twins based on BIM:

- *Better understanding:* the Digital Twin allows you to get a more complete understanding of the building and its functioning.
- Effective management: the Digital Twin helps to more effectively manage the building throughout its life cycle.
- *Optimization*: the Digital Twin allows you to optimize various aspects of building operation, such as energy consumption, comfort, and safety.
- *Forecasting:* the Digital Twin allows you to predict the behavior of the building in different conditions and scenarios.

In general, BIM and Digital Twins are interconnected technologies that complement each other. BIM models are the basis for creating Digital Twins, and

Digital Twins, in turn, expand the capabilities of BIM models, allowing modeling the behavior of buildings in dynamics and solving various tasks.

A Digital Twin consists of three main components:

- *Physical object*: the real object for which the Digital model is created.
- *Virtual model*: a digital copy of a physical object that contains information about its characteristics and behavior.
- *Connection between them*: provides data exchange between the physical object and its virtual model.

Different components of Digital Twins can be provided by different manufacturers and suppliers. Although they should work together, at least in theory, in practice this is not the case. Moreover, artificial intelligence (AI) and modeling tools, which are expected to perform the same function, may not support this capability. It is not uncommon to find an AI tool or modeling application from one supplier that is not able to replicate the capabilities of another supplier's product. This creates an *interoperability problem*. Different building blocks of Digital Twins form a large attack surface that cybercriminals can target. And given the need for hundreds and thousands of sensors connected to the Internet, the attack surface becomes even larger. Of course, the intentions of the attackers may be due to the importance of Digital Twins and related data for the activities of organizations. Having gained access to service data, criminals can demand large sums of cash from companies so that they do not disclose the illegally obtained find or compromise the Digital Twin.

Various undesirable consequences of security breaches indicate the need to give priority to *cybersecurity*. But effectively managing the cybersecurity of Digital Twins can be a daunting task for some organizations, given the vast amount of software, parts, and specialists needed to create a working digital twin.

Digital Twins are an important tool for increasing efficiency and optimizing various processes. They allow [10]:

- Predict the behavior of objects and systems.
- Model different scenarios and make informed decisions
- Optimize production processes and reduce costs.
- Improve the quality of products and services.
- Create new products and services.

In general, a Digital Twin acts as a single source of information about a project, which helps to improve collaboration. In addition, it provides all stakeholders with a deeper understanding of the products, processes, environments, and personnel involved in the project. It is also worth noting that several Digital Twins can be integrated, providing a deeper understanding of the interdependencies and the ecosystem in which they exist.

Digital Twins are classified according to different criteria, depending on the purpose and scope of application. Here are some of the most common types:

By level of detail:

- *Digital Twins of objects*: reflect individual physical objects, such as an engine, a machine tool, or a building.
- *Digital Twins of processes:* model technological or business processes, for example, production of products or logistics.
- *Digital Twins of systems:* describe complex systems consisting of many components, such as the energy system of a city or a transport network.

By purpose:

- *Digital Twins for design:* used to develop and design new products or systems.
- *Digital Twins for production:* used to manage production processes, optimize equipment operation, and control product quality.
- *Digital Twins for operation*: serve to monitor the condition of equipment, predict breakdowns and maintenance.

By the nature of the connection with the real object:

- *Digital Twins that reflect the past*: contain information about the object for a certain period of time in the past.
- *Digital Twins that reflect the present*: reflect the current state of the object in real time.
- *Digital Twins that predict the future:* used to predict the behavior of an object in the future based on data analysis and modeling.

By complexity:

- *Simple Digital Twins:* contain a limited amount of information about the object and its behavior.
- *Complex Digital Twins:* include detailed information about the object, its interaction with the environment, and complex behavior models.

One Digital Twin can combine several types, for example, be both a Digital Twin of an object that reflects the present and is used to predict the future. The choice of the type of Digital Twin depends on the specific task and requirements for the accuracy of the model. It is important to note that the development of technologies, such as artificial intelligence and machine learning, contributes to the creation of increasingly complex and functional Digital Twins, which find applications in various industries.

APPLICATIONS OF DIGITAL TWINS IN CAD

Digital Twins are used in many areas, helping to solve various problems and optimize processes. Here are some specific examples [7,10]:

Industry

- *Manufacturing:* Digital Twins are used to model production lines, optimize equipment operation, predict breakdowns, and manage product quality. For example, General Electric uses Digital Twins of aircraft engines to monitor their condition in real time and predict the need for maintenance.
- *Energy*: Digital Twins help in managing energy systems, forecasting consumption, optimizing equipment operation, and integrating renewable energy sources.
- *Automotive*: Digital Twins are used to design cars, simulate their behavior on the road, optimize production, and develop driver assistance systems.

Cities and infrastructure

- *Smart cities:* Digital Twins of cities are used to manage traffic flows, optimize the operation of utilities, monitor the state of infrastructure, and respond to emergencies.
- *Construction:* Digital Twins of buildings help in the design, construction, and operation of buildings, allowing to optimize the use of resources, control the quality of work, and predict the condition of structures.

Healthcare

- *Personalized medicine*: Digital Twins of patients can be used to simulate the body's response to various treatments and predict the development of diseases.
- *Drug development:* Digital Twins help in the development of new drugs by simulating their effect on the body and predicting effectiveness.

Other areas

- *Logistics:* Digital Twins help in optimizing delivery routes, managing warehouse stocks, and forecasting demand.
- *Agriculture:* Digital Twins of fields are used to monitor the condition of crops, predict yields, and optimize the use of resources.

These are just some examples of the use of Digital Twins. With the development of technology, the scope of their use is constantly expanding, opening up new opportunities for optimizing and managing various processes. *Usually, CAD is a fundamental step towards creating a Digital Twin; it is the foundation* [11]. Therefore, *it can be said that without CAD there is no Digital Twin*. In fact, Digital Twin platforms integrate with CAD and BIM solutions. For example, Autodesk Tandem, a Digital Twin platform, is designed to integrate CAD geometry with Revit, geospatial data, facility management data, IoT data.

It is possible to develop a "fault dictionary" for the future Digital Twin using mathematical modelling and CAD tools during the design of objects and processes. To do this, it is necessary to identify in advance a set of influential parameters or model variables to measure which control ports should be provided in the structures and processes to be designed. Then, using powerful techniques for multivariate analysis, sensitivity analysis, optimization, worst-case evaluation of the deviation of component parameters from nominal values (WCD), and the inverse of this assigning optimal tolerances, statistics can be collected for DC, TR, and AC modes on the deviation of the output parameters of the object (process) from the effects of destabilizing factors (temperature, radiation, humidity, etc.) and, for example, component "aging" over time. All this can be amplified by statistical analysis, which introduces certain distribution laws for the above-mentioned deviations of component parameters. The availability of such a pre-modelled "fault dictionary" will greatly facilitate the interpretation of measurement data taken on a real installation manufactured by the design documentation. If the mathematical and CAD models are kept together with the modelled "fault dictionary", the transition to the Digital Twin will be reduced to additional investigations of these models, taking into account the measured data from the real object, if the "fault dictionary" information is not sufficient.

THE PROCESS OF CREATING A DIGITAL TWIN

The process of creating a Digital Twin can be imagined as a sequence of steps, each of which has its own characteristics and requires certain resources.

1. **Defining the goal and scope:**

- At this stage, it is important to clearly define why a Digital Twin is being created (monitoring, analysis, simulation, optimization, etc.), what tasks it should solve, what parameters of the object or process need to be modeled.
- It is also important to determine what level of detail of the model is needed to achieve the goals (machine, production process, building, city). Today,

it is common to refer to objects with a deep interconnection between their physical and computational elements as "*cyber-physical systems*" [12,13].

2. **Data collection:**

- To create a Digital Twin, data about the object or process being modeled is required.
- This can be data on the physical properties of the object (geometry, materials, physical characteristics), sensor data on the state of the object (temperature, vibration, pressure, etc.), previous data on the operation of the object, accidents, maintenance.
- It is important to ensure the quality and reliability of the data, since the accuracy and adequacy of the Digital Twin depends on this.

3. **Building a model through modeling and simulation:**

- *3D Modeling:* creating a three-dimensional model of an object using CAD (software).
- *Physical models:* development of models that describe physical processes (mechanical, thermodynamic, electromagnetic models).
- *Behavioral models*: models that predict the behavior of the system under various conditions.

4. **Data integration:**

- *Platform*: selection or development of a platform that can integrate all data streams and ensure interaction between physical and digital objects.
- *API and Interfaces:* development or use of existing APIs for data exchange between different systems.

5. **Analytics and Data Processing**

- *Analytical tools:* using machine learning, artificial intelligence to analyze data, predict wear and tear, optimize work.
- *Visualization:* creating interfaces for visualizing the status and results of the analysis.

6. **Validation and verification:**

- *Testing:* checking the accuracy of the Digital Twin by comparing its behavior with a real object.
- *Adjustment:* Making changes to the model based on real data to improve accuracy.

7. **Deployment and use:**

- *Integration into business processes*: using a Digital Twin in daily operations, maintenance planning, equipment modernization.
- *Personnel training:* training users to work with a Digital Twin.

8. **Update and support:**

- *Monitoring:* continuous monitoring of both the physical object and the Digital Twin to detect deviations.
- *Updating:* regularly updating the model with new data to maintain relevance.

Creating a Digital Twin is a complex and multi-stage process that requires significant resources and competencies. However, a properly created Digital Twin can be a powerful tool for increasing efficiency, optimizing processes, and making informed decisions.

Recall that digital modeling is the basis for the functioning of Digital Twins, so the choice of modeling algorithms plays a dominant role in their implementation due to the significant difference in the nature of objects and the tasks they perform [11]. This can be demonstrated by examples of specifications of Digital Twins for a rocket flight control unit, for a bridge structure and a process microcontroller (PCM), presented in Fig. 1, Fig. 2 and Fig. 3, respectively.

DIGITAL TWIN OF A ROCKET FLIGHT CONTROL UNIT

Main tasks:

- Specifications of the Digital Twin of the rocket flight control unit.
- Modeling the dynamics of rocket flight in real time.
- Predicting the trajectory and possible deviations.
- Controlling the guidance system and trajectory correction.
- Diagnostics and troubleshooting.

Algorithms and methods:

Mathematical modeling:

- Differential equations of motion (Newton’s, Euler’s laws) taking

into account interaction with

- the atmosphere and space environment.
- Models of aerodynamics and jet propulsion.
- Trajectory calculation algorithms and ballistic models.

Signal and data processing:

- Kalman filters for processing data from sensors.
- Methods of statistical analysis to estimate flight parameters.
- Pattern recognition algorithms to detect deviations.

Artificial intelligence and machine learning:

- Neural networks for forecasting and control.
- Reinforcement learning algorithms for control optimization.

Fig. 1. Specifications of the Digital Twin of the rocket flight control unit

DIGITAL TWIN OF A BRIDGE STRUCTURE

Main tasks:

- Modeling the behavior of the bridge under load.
- Analysis of the strength and stability of the structure.
- Predicting wear and possible damage.
- Monitoring the condition of the bridge in real time.

Mathematical modeling:

Finite element method (*FEM*):

- Dividing the structure into many elements.
- Calculation of stresses and strains in each element.
- Analysis of the general behavior of the structure under load.

Material modeling:

- Models of elasticity, plasticity and fracture of materials.
- Taking into account the influence of temperature, humidity and other factors.

Statistical analysis:

- Risk assessment and probability of damage.
- Predicting the residual life of the structure.

Artificial intelligence and Machine learning:

- Classification algorithms for detecting defects based on data.
- Neural networks for predicting durability.

Fig. 2. Specifications of the digital twin of a bridge structure

DIGITAL TWIN FOR A PROCESS MICROCONTROLLER (PCM)

Main tasks:

- *Real-time modelling*: the Digital Twin should reflect the state of the PCM with high accuracy and synchronization.
- *Interaction with the external environment*: the digital twin should take into account the impact of external factors on the operation of the PCM
- *Debugging and testing capability:* the Digital Twin allows developers to test the operation of the IPCM without using additional equipment.

Modelling the logic of work:

- *Boolean algebra*: to describe logical operations, input/output states, switching conditions.
- *State machine*: to simulate the sequence of actions, transitions between operating modes.
- Controller programming languages (Ladder Logic, ST, FBD): to describe process control algorithms.

Modelling process dynamics:

- *Differential equations:* to describe the dynamics of physical quantities (temperature, pressure, level, etc.).
- *Difference equations*: for discrete modeling of processes in time.
- *Transfer functions*: to describe the dynamic properties of control objects.

Signal processing algorithms:

- *Filtering*: to clean signals from noise and distortion (Kalman filter, moving average).
- *Conversion*: to convert signals from analog to digital form and vice versa (ADC, DAC).
- *Scaling*: to bring signals to the desired range of values.

Communication algorithms:

- Data exchange protocols (Modbus, Profibus, Ethernet/IP): to ensure communication between the PCM and other systems.
- Synchronization algorithms: to coordinate the operation of various components of the digital twin.
- Error handling mechanisms: to ensure communication reliability.

Visualization and interface algorithms:

- *Graphic libraries*: to create a visual representation of the state of the PCM and the technological process.
- *Data visualization algorithms*: to display trends, graphs, charts.
- *User interfaces*: to provide convenient access to information and control.

Fig. 3. Specifications of the Digital Twin of a process microcontroller

Differences and features

- *Different physical models*: for a rocket models of flight dynamics and aerodynamics are used, for a bridge — models of structural mechanics and resistance of materials, for a process microcontroller — models of functioning under the influence of external factors.
- *Different time scales:* rocket flight and microcontroller reactions occur quickly, so real-time modeling is required. The bridge is operated for a long time, so long-term forecasting is important.
- *Different types of data:* for a rocket, data from sensors and control systems are used, for a bridge — data on loads, the state of materials and structures, for a microcontroller — data from sensors about the state of the technological process.

That is, the creation of digital twins for different objects requires the use of different algorithms and modeling methods, due to the specifics of the tasks and physical principles underlying the functioning of each of them.

CONCLUSIONS

Digital Twins are becoming increasingly commonplace in many industries, shaping the future. The growing adoption of Digital Twin technology is driven by the convergence and development of technologies such as the Internet of Things, sensors, artificial intelligence, machine learning, cloud computing, simulations, and more. This technology uses data to bring to life once static CAD models. This gives many advantages.

Digital Twins can describe physical objects, diagnose problems, analyze results, and predict future events. Due to these advantages, companies and organizations in the manufacturing, medical, mining, infrastructure and planning industries, agriculture and logistics are implementing Digital Twins. However, their implementation has not been without certain problems. From low data quality and lack of data standardization to complex change management, cybersecurity threats, inaccurate representation, and the need to protect intellectual property. Fortunately, there are ways to get around these problems.

Let's present a comparison of the basic properties of CAD, BIM and Digital Twin technologies in the form of a table.

Comprehensive comparison between CAD, BIM and Digital Twins

Feature	CAD (Computer-Aided Design)	BIM (Building Information Modeling)	Digital Twins
Main function	Creating 2D and 3D models of for objects	Creating and managing building information throughout its lifecycle	Virtual representation of a physical object or process, reflecting its behavior in real time
Model type	Geometric model	Information model	Dynamic model
Data	Geometry, dimensions, materials	Geometry, dimensions, materials, structures, systems, equipment, lifecycle data	Data from sensors, status information, historical data, forecasts
Connection with the real object	Static	Static with dynamic elements	Dynamic, updated in real time
Applications	Design, visualization, documentation	Design, construction, operation of buildings	Monitoring, analysis, optimization, forecasting, training
Examples	Drawings of parts, models of machines, diagrams	3D model of a building with information about all its elements	Digital Twin of an aircraft engine, virtual copy of a city, patient model
Limitations	Limited information about the object, static model	Limited information about the dynamics of the object	Complexity of creation and maintenance, requires a large amount of data

Key differences:

- CAD focuses on the geometry and shape of the object, while BIM includes a wider range of information about the building, including its functionality and characteristics.

- Digital Twins go even further, providing a dynamic representation of an object or process that is updated in real time.
- CAD and BIM are mainly used for design and construction, while Digital Twins have a wider range of applications, including monitoring, analysis, and optimization.

Artificial intelligence (AI) is significantly impacting the development of Digital Twins, enhancing their capabilities and expanding their scope of application. Here are some key aspects of this impact [14]:

- **Real-time data analysis**

Prediction: AI can analyze data streams coming from IoT sensors to predict the future state or behavior of a physical object. For example, AI can predict when a machine will need maintenance before a breakdown occurs.

Optimization: Using machine learning, Digital Twins can optimize the operation of systems, such as the energy consumption of buildings or the performance of production lines.

- **Process automation**

Automatic model updates: AI can automatically update digital models based on new data, ensuring that the Digital Twin is up-to-date with its physical counterpart.

Automation of response: AI can not only detect anomalies or problems, but also automatically initiate appropriate actions, such as ordering spare parts or adjusting system parameters.

- **Improving simulations**

Complex scenarios: AI allows modeling and testing much more complex and diverse scenarios, making it possible to explore "what if" situations without risk to the real object.

Adaptive learning: AI models can learn from historical data and real-world interactions, improving the accuracy of simulations over time.

- Personalization and adaptation

Individual solutions: AI can adapt Digital Twins for individual needs, such as customizing the user interface in cars or adapting medical equipment for a specific patient.

- **Improving human-machine interaction**

Natural language and interfaces: AI, especially NLP (natural language processing), allows more natural interaction with Digital Twins through voice commands or text queries.

Virtual assistants: Using AI to create virtual assistants that can help manage Digital Twins by providing recommendations or performing tasks.

- **Security and cybersecurity**

Threat detection: AI can be used to detect anomalies that may indicate cyberattacks or other security problems in systems associated with Digital Twins.

- **Future development**

Cognitive Digital Twins: In the future, the development of cognitive Digital Twins is expected, which not only reflect the state of the object, but can also "think" and "make decisions" similar to their physical counterpart, using advanced AI algorithms.

Integration with blockchain: To ensure data security and trust in Digital Twins, blockchain technology can be used to ensure data immutability and transparency.

Thus, AI not only improves the capabilities of Digital Twins, but also expands their application, making them more intelligent, adaptive, and integrated into various aspects of human life and industry.

REFERENCES

1. A. Petrenko, "Optimal Engineering Design Based on Service-Oriented Computing (SOC). Chapter 7," *Book "Research Highlights in Science and Technology,"* Vol. 4. UK: B P International publishes, 17 June 2023, pp. 97–106. doi: https://doi.org/10.9734/bpi/rhst/v4/10278F
2. A. Petrenko, V. Ladogubetz, V. Chkavov, J. Pudlowski, *ALLTED: a Computer-Aided Engineering System for Electronic Circuit Design*. Melbourne, UICEE, 1997, 206 p. Available: https://unesdoc.unesco.org/ark:/48223/pf0000156798
3. *CAD and BIM: Bridging the Gap Between Digital and Physical Construction - Current Trends and Insights*. Available: https://www.linkedin.com/pulse/cad-bim-bridging-gap-between-digital/
4. *BIM vs CAD: Exploring the key differences*. Available: https://www.bimcollab.com/en/resources/ blog/bim-vs-cad-comparison/
5. Liu Zihe, "Comparison and analysis of advantages and disadvantages between BIM and CAD in civil drafting software," *Applied and Computational Engineering*, 62(1), pp. 192–197, May 2024. doi: 10.54254/2755-2721/62/20240426.
6. *Scan to BIM: Reality capture meets actionable data models*. Available: https://www.autodesk.com/industry/land-development/scan-to-bim
7. N.D. Pankratova, K.D. Grishyn, V.E. Barilko, "Digital Twins: stages of concept development, areas of use, prospects," *System Research and Information Technologies*, no. 2, pp.7–21, 2023. doi: https://doi.org/10.20535/SRIT.2308-8893.2023.2.01
8. F. Tao, H. Zhang, A. Liu, A.Y.C. Nee, "Digital Twin in Industry: State-of-the-Art," in *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2405–2415, April 2019. doi: 10.1109/TII.2018.2873186.
9. *Digital Twin Technology & CAD, February 2, 2025, Computer-Aided Engineering*. Available: https://cadifystudio.com/digital-twin-technology-and-cad-its-benefits-and-applications/
10. Angira Sharma, Edward Elson Kosasih, Jie Zhang, Alexandra Melike Brintrup, "Digital Twins: State of the art theory and practice, challenges, and open research questions," *Journal of Industrial Information Integration*, 30(1), 100383, August 2022. doi: 10.1016/j.jii.2022.100383.
11. L. Wright, S. Davidson, "How to tell the difference between a model and a Digital Twin," *Adv. Model. and Simul. in Eng. Sci.*, 7, 13 (2020). doi: https://doi.org/10.1186/s40323-020-00147-4
12. Nataliya Pankratova, Igor Golinko, "Electric heater mathematical model for cyber-physical systems," *System Research and Information Technologies*, no. 2, pp. 7–17, 2021. doi: 10.20535/SRIT.2308-8893.2021.2.01.
13. Nataliya Pankratova, Igor Golinko, "Development of Digital Twins to Support the Functioning of Cyber-physical Systems," *Computer Science Journal of Moldova*, vol. 31, no. 3(93), pp. 299–320, 2023. doi: https://doi.org/10.56415/csjm.v31.15
14. Tim Kreuzer, Panagiotis Papapetrou, Jelena Zdravkovic, "Artificial intelligence in Digital Twins—A systematic literature review," *Data & Knowledge Engineering*, vol. 151, May 2024. doi: https://doi.org/10.1016/j.datak.2024.102304

Received 20.02.2025

INFORMATION ON THE ARTICLE

Anatolii I. Petrenko, ORCID: 0000-0001-6712-7792, Educational and Research Institute for Applied System Analysis of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine, e-mail: tolja.petrenko@gmail.com

ВІД CAD і BIM ДО ЦИФРОВИХ ДВІЙНИКІВ / А.І. Петренко

Анотація. Цифровий двійник — це віртуальна модель фізичного об'єкта або системи, яка використовує дані в реальному часі для моделювання поведінки свого реального аналога. Це може бути продукт, машина, будівля або навіть ціле місто. Цифрове моделювання є фундаментальною технологією для створення цифрових двійників, оскільки воно забезпечує методологію для подання фізичних об'єктів у віртуальному світі. Computer-Aided Design (CAD) може бути використано для створення початкової моделі цифрового двійника. Building Information Modeling (BIM) є спеціалізованою формою CAD, яка зосереджена на будівельних проектах і містить більше інформації, ніж просто геометрія. BIM включає в себе не тільки геометрію, але й дані про час, витрати, експлуатацію та обслуговування. Досліджено, як ці технології дедалі більше інтегруються одна в одну, базуючись на застосуванні математичного моделювання, яке шляхом проведення віртуальних обчислювальних експериментів забезпечує розуміння складного функціонування об'єктів та прийняття обґрунтованих рішень у різних сферах. Інтеграція цифрових двійників і CAD трансформує способи проектування, моделювання та оптимізації продукції в промисловості. BIM моделі можуть стати основою для створення цифрових двійників будівель, які потім використовуються для оптимізації енергоспоживання, обслуговування та ремонту. З ростом інтернету речей (IoT), цифрові двійники отримують дедалі більше реальних даних, що робить їх ще більш точними і корисними для прогнозування та оптимізації. Використання штучного інтелекту для аналізу даних, зібраних цифровими двійниками, дозволяє прогнозувати поломки, оптимізувати процеси і навіть автоматизувати управління системами.

Ключові слова: математичне моделювання, цифрове моделювання, CAD і BIM, цифрові двійники (ЦД), інтернет речей (IoT), застосування штучного інтелекту в ЦД.

UDC 004.4
DOI: 10.20535/SRIT.2308-8893.2025.2.02

STREAMLINED MANAGEMENT OF PHYSICAL AND CLOUD INFRASTRUCTURE THROUGH A CENTRALIZED WEB INTERFACE

I. BYZOV, S. YAKOVLEV

Abstract. The article focuses on the development of a centralized management system for both physical and virtual servers via a web interface. This system enables server infrastructure administration through various tools. Key components of the system, integration with server management tools, the use of Paramiko for remote access via SSH, and the potential application of other technologies for managing cloud and physical servers through a web interface are discussed. The system enables centralized management of various server types through a web interface, supporting both basic and advanced administrative operations. Integration with Docker and support for cloud APIs provide convenient automation and simplify operations with both cloud and physical servers. As a result, the system serves as a versatile tool for DevOps engineers. The study highlights the relevance and importance of a centralized approach to server infrastructure management, thereby enhancing the efficiency and reliability of server operations in modern environments.

Keywords: centralized management, SSH, web interface, cloud services, automation, DevOps.

INTRODUCTION

With the rapid growth of information technology and the increasing volume of data processed by server infrastructures, effective management of physical and virtual servers has become a critical necessity. Hybrid infrastructures, which combine physical servers and cloud computing resources, are widely adopted for their flexibility and scalability in handling data processing and storage. However, the expansion of these infrastructures introduces significant complexity into administrative tasks, demanding more advanced and streamlined management solutions.

Server administrators frequently encounter challenges in managing diverse server types—both physical and virtual—across multiple tools and protocols. This fragmentation leads to inefficiencies, as routine tasks such as updates, monitoring, configuration, and troubleshooting must be performed using disparate platforms and interfaces.

This study addresses these challenges by presenting a centralized management system for physical and virtual servers via a web interface. The proposed solution aims to automate server administration processes by integrating various modern tools and technologies. Specifically, it incorporates Docker for containerization, Paramiko for secure SSH-based remote access, and cloud APIs for managing virtual servers across different providers. This unified approach not only optimizes administrative workflows but also minimizes errors and improves overall efficiency.

The primary goal of this research is to develop a universal solution for automating the management of hybrid server infrastructures, including both physical servers and cloud instances. The proposed system offers a centralized web interface for managing diverse server environments, integrating technologies such as Docker, Terraform, Ansible, and cloud APIs. This integration simplifies routine administrative tasks, enhances productivity, and reduces complexity and error rates associated with fragmented tools.

The article is structured as follows. The next section provides a detailed review of existing tools and technologies for managing physical and virtual servers, with particular emphasis on solutions such as Docker, Paramiko, and cloud APIs. Section 3 formulates the main research objectives. Section 4 describes the architecture of the proposed centralized management system, highlighting its design and the integration of various server management tools within a unified web interface.Subsequent sections focus on the system's implementation, detailing the integration of Docker, Terraform, Ansible, and Paramiko into a cohesive solution. A comparison with existing management tools is also included.

Finally, the effectiveness of the developed system is evaluated through performance testing and automation analysis. The conclusion summarizes the findings and discusses potential avenues for further development and enhancement of the system.

CURRENT STATE OF THE PROBLEM

Various scientific studies highlight the importance of automation and centralization in server infrastructure management. One of the primary trends in this field is the use of containerization to ensure efficient resource allocation and process isolation. Technologies such as Docker and Kubernetes are among the most prominent in this area. Docker facilitates the containerization of software applications, while Kubernetes orchestrates container operations, enabling the automation of deployment, scaling, and management processes.

In [1], the authors showed that Docker significantly simplifies infrastructure management through application isolation and streamlined deployment processes. This fact is confirmed by studies [2] emphasize Kubernetes' advantages, particularly in cloud environments. Kubernetes offers a high degree of automation, allowing efficient resource management and infrastructure scaling while also highlighting opportunities to enhance Kubernetes (K8s) for deployment platforms.

Another critical aspect of server management is the use of the SSH protocol for remote administration of physical servers. Many automation tools, such as Ansible, Chef, and libraries like Paramiko, leverage SSH to execute commands on remote servers. Among these, Paramiko has been identified as one of the most efficient and user-friendly tools. A study [3] showed that Paramiko reduces server configuration time by 4.14 times compared to Netmiko.

On the other hand, numerous researchers focus on managing cloud resources via APIs. For example, in [4], explores methods of integrating with cloud services such as AWS, Azure, and Google Cloud. These integrations enable automation in managing cloud instances, significantly reducing administrative overhead.

Most studies conclude that centralizing hybrid infrastructure management is an effective approach to reducing the complexity of administration and improving server management productivity. In [5] it is emphasized that unifying tools for cloud infrastructure management into a single interface greatly simplifies administrative processes, reduces error rates, and enhances overall productivity.

A review of the literature underscores the critical relevance of centralized management for hybrid infrastructures in modern IT systems. Existing research emphasizes the need for solutions that integrate essential management tools while modernizing algorithms to enable efficient centralized administration. These approaches simplify infrastructure management and enhance overall server operation efficiency, marking an important direction for further research and implementation in the DevOps domain.

RESEARCH OBJECTIVES AND TASKS

Managing server infrastructures, both physical and cloud-based, has become increasingly complex and demanding in modern IT environments. Numerous tools exist to address specific tasks, such as containerization, remote server access, and cloud resource management. However, effective management of hybrid infrastructures requires a centralized system that integrates all these tools into a single interface, simplifying administration and automating routine tasks.

The primary objective of this research is to design the architecture and develop a web application that provides a unified interface for managing physical and cloud servers, integrating key tools and technologies.

To achieve this objective, the following tasks are defined:

- *Designing the architecture of a centralized management system.* Develop the architecture of a web application capable of integrating various tools for managing physical and cloud servers within a unified interface.
- *Implementing integration with Docker, SSH, and cloud services.* Enable interaction with Docker for containerization, utilize Paramiko for secure remote access to physical servers via SSH, and integrate cloud services through APIs.
- *Testing and evaluating the system's performance.* Conduct experimental testing of the developed system on various infrastructures, including physical servers and cloud platforms, to assess its efficiency and performance in real-world scenarios.
- *Analyzing results and identifying future development prospects.* Based on the testing results, draw conclusions about the feasibility of implementing the system in production environments and identify potential directions for further development and enhancement.

This structured approach ensures the development of an efficient, centralized solution that simplifies hybrid infrastructure management and optimizes server administration processes.

METHODS AND TOOLS FOR INFRASTRUCTURE MANAGEMENT

Automation of server tasks through specialized modules using the SSH protocol significantly reduces the time required for routine tasks, such as installing updates, restarting services, scheduling tasks, or managing Docker containers. The proposed system also facilitates rapid infrastructure scaling by enabling server configuration with IaC tools like Ansible and Terraform via preconfigured templates and scripts accessible through a web interface. This approach minimizes human error and improves server reliability.

Integration with cloud provider APIs adds another layer of convenience, allowing real-time monitoring and management of cloud resources. Users can fetch data from multiple providers into a single web application and perform necessary operations across providers through a unified platform.

The developed system thus provides an efficient solution for centralized management of large-scale server infrastructures, reducing maintenance costs and improving the productivity of administrators and DevOps engineers.

Tools for Infrastructure Management (Ansible, Terraform, Docker)

Modern automation and infrastructure management tools, such as Ansible, Terraform, and Docker, play a crucial role in establishing and maintaining effective DevOps processes. Ansible is a powerful tool for configuration management and automation, supporting both physical and virtual servers. Its modular architecture and user-friendly interface make it ideal for scaling infrastructure management processes, automating repetitive tasks, and reducing the risk of errors.

Terraform specializes in Infrastructure as Code (IaC), enabling administrators to automate the provisioning and management of infrastructure using declarative configurations. Its support for multiple cloud providers and modules makes it particularly effective for building cloud infrastructure from scratch using codified scripts.

Docker provides lightweight containerization for isolating applications on servers. It allows multiple applications or services to run on a single server in isolated environments, enhancing resource utilization and simplifying deployment and application management. Integrating Docker with the web-based management system adds a new level of automation and flexibility for both physical and cloud environments.

By integrating these tools into a centralized management system with an intuitive web interface, the proposed solution enhances the efficiency of administration and automation processes, making it a vital asset for modern IT operations.

System Architecture

The architecture of the proposed system is based on a web application developed in Python using the Django framework, which enables centralized management of hybrid infrastructures. The system's modular structure provides flexibility for easy expansion, allowing new features to be integrated, such as support for additional cloud providers or other infrastructure management tools.

At its core, the architecture includes a server-side component that handles SSH requests via the Paramiko library, integration with Docker for managing containers on remote servers, and API interactions with various cloud providers for managing their infrastructure. The key components of the system are the backend server, the web interface for user interaction, and authentication and authorization mechanisms to ensure security.

By utilizing the paramiko.SSHClient() function and intermediate servers, the web application can execute all the functions shown in Figs. 1, 2, enabling seamless interaction and operation management.

To ensure that all functions can be performed, several variations of sending commands to the server have been developed, including:

- d connection through the terminal to the user's server;
- sending a command via a request with a response.

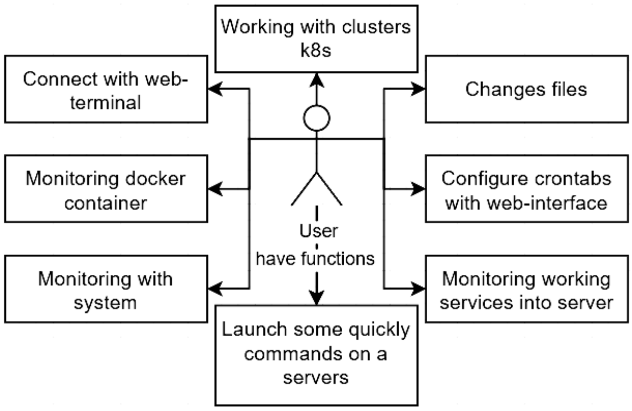


Fig. 1. User options

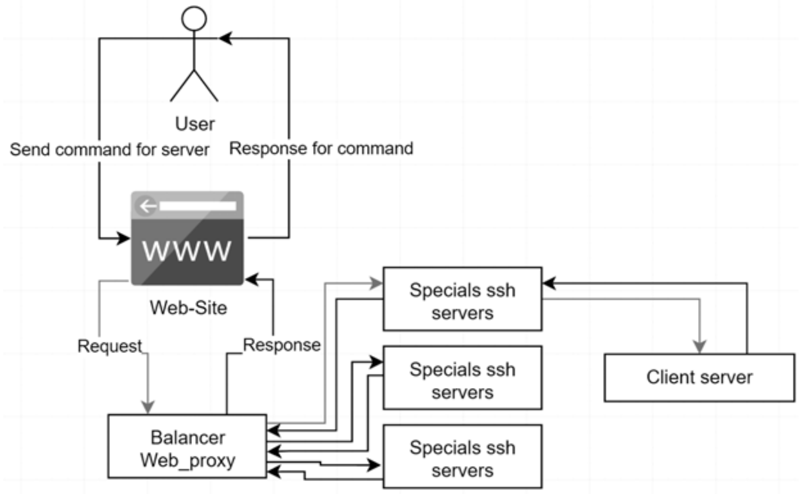


Fig. 2. Scheme of user and application interaction

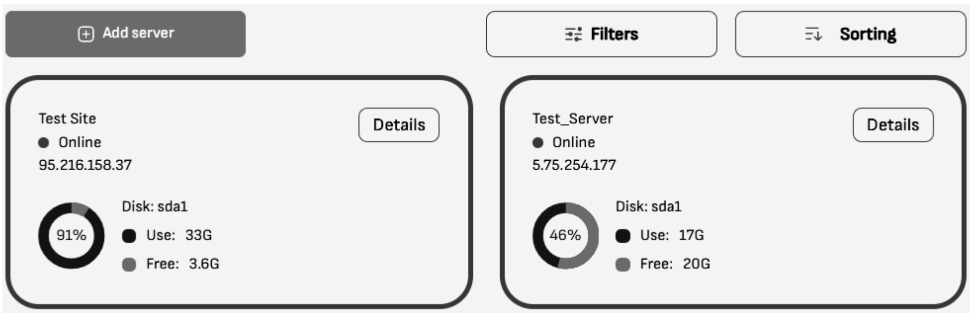


Fig. 3. Servers display page

The primary principle underlying this approach is that users can perform operations on their own servers from other servers, even in conditions where power outages periodically occur in Ukraine and the quality of the Internet connection can be unstable. The utilization of Paramiko for remote server access through a web application offers several key advantages. In the event of an unreliable connection, a request-response cycle is employed, and establishing a connection via a terminal will also function correctly, as it does not necessitate a full connection to the target server.The application can both automatically receive data about cloud servers, having received all the necessary information from the provider's API, and provide the ability to add a separate server to the site, which can be either physical or another cloud provider, which allows you to quickly switch and start working with it (Fig. 3).

Automation of server operations

Secure Shell (SSH) is a standard protocol for secure server access that plays a crucial role in automating server infrastructure management processes. The utilization of SSH facilitates remote server management, command execution, service management, and automation of diverse operations, obviating the need for physical access to the servers. The ability to execute commands via terminal and transfer files renders SSH a versatile tool for managing all types of servers [7].

In contemporary environments, the implementation of SSH for the automation of operations substantially enhances the flexibility and speed of administration, particularly in distributed infrastructures where servers are dispersed across multiple geographical regions. Integrating the SSH protocol into a website enables administrators to perform operations on remote servers via terminal, thereby simplifying management and automation processes.

The utilization of SSH as a transport mechanism by prominent automation tools such as Ansible and Chef underscores its pivotal role in contemporary infrastructure automation solutions. Another key aspect is the security provided by SSH through encrypted connections and key-based authentication [8].

Tools for Centralized Management of Cloud and Physical Servers

In the contemporary landscape, numerous solutions have emerged that facilitate the centralized management of both cloud and physical servers. A prominent example is Mirohost, and Red Hat Satellite is an infrastructure management solution designed to administer and monitor servers based on Red Hat Enterprise Linux (RHEL) and other Red Hat products. The MiroHost website facilitates the management of physical and virtual servers within a centralized system, though this functionality is exclusive to Mirohost's servers (Figs. 4, 5, 6) [9].

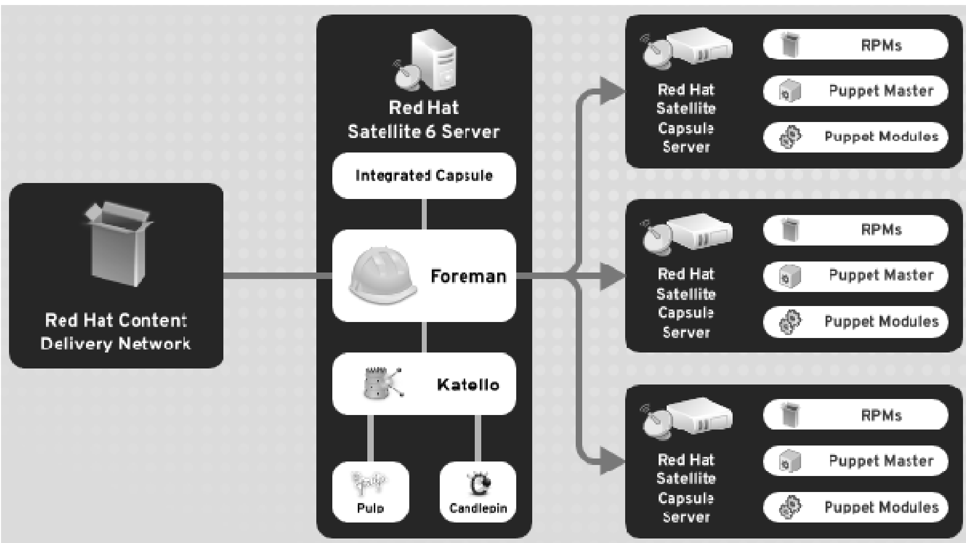


Fig. 4. Red Hat Satellite interaction diagram

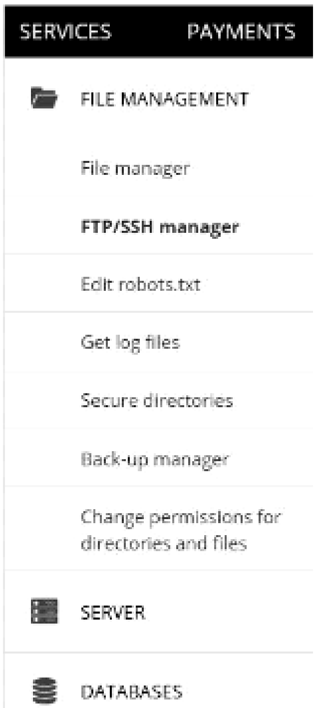


Fig. 5. Admin panel for Mirohost interaction

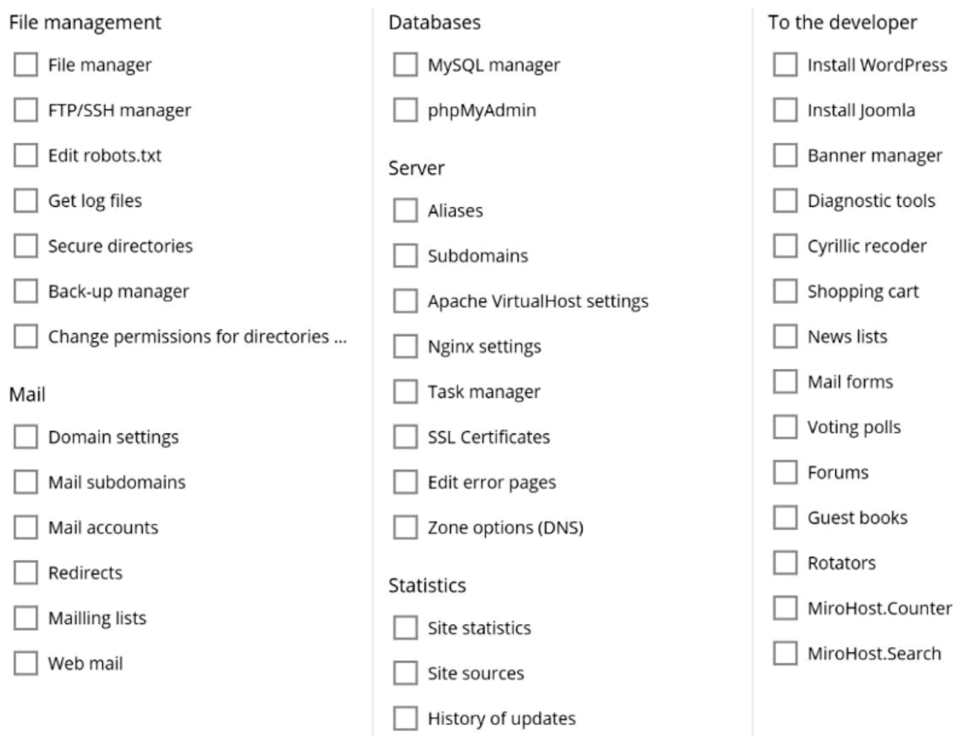


Fig. 6. Mirohost functions for interacting with the server

However, the majority of existing web-based applications lack seamless integration for concurrent management of physical and cloud resources through a unified web interface. Other applications can be cumbersome to utilize or necessitate specialized configuration for various providers and servers.The study demonstrates that these limitations can be addressed by integrating the management capabilities of both physical and virtual servers within a single system. This approach has the potential to enhance the processes of backup, monitoring, and updating the infrastructure. The system of operation of the web application is very similar to the system described in (Fig. 7) [10–12].

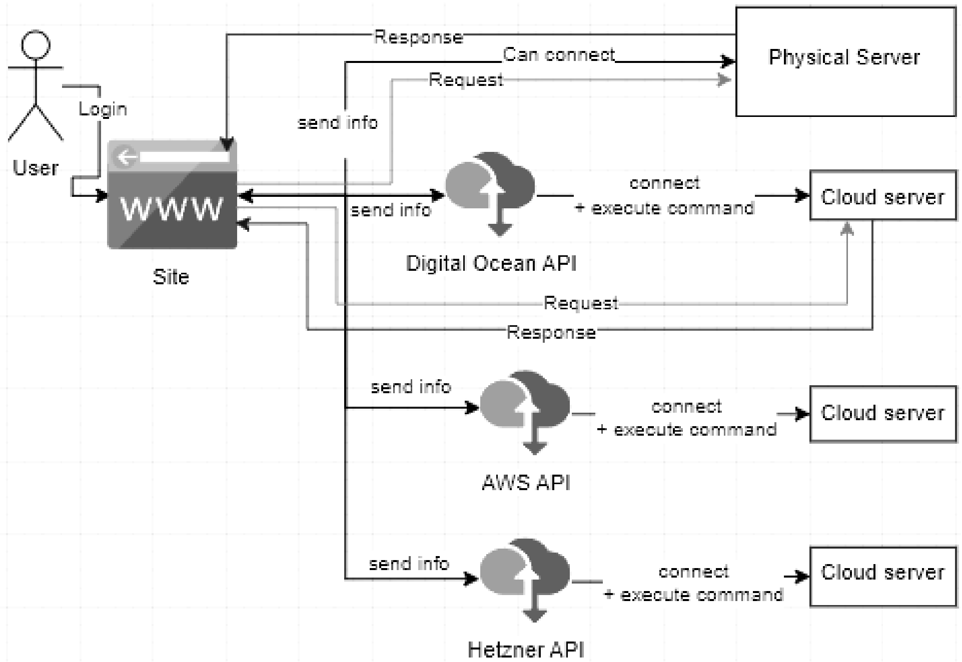


Fig. 7. Scheme of user interaction with the application

Use of modern technologies and tools

A pivotal technology in the implementation of the system is the Paramiko library, which facilitates Secure Shell (SSH) connections to all servers [13; 14]. The utilization of SSH ensures flexibility and security in the management of disparate server types, irrespective of their geographical location or categorization.

The system incorporates a user data encryption mechanism, ensuring that added passwords and keys are encrypted using a unique algorithm. This feature facilitates secure connections to servers via the terminal and enables the execution of commands remotely through the web interface, thereby eliminating the necessity for direct interaction with the terminal.The Paramiko library plays a pivotal role in establishing reliable and secure connections to servers, as well as in the transmission of commands to administrators. This includes the management of system services and the configuration files of added servers.

The system also includes the ability to manage Docker containers (Fig. 8), which allows users to start, stop, and view the status of containers through a web interface. This is achieved with the help of Paramiko and additional modules integrated into the server side.

This module enables specialists to efficiently manage containers on remote servers. The integration of this module significantly simplifies the process of managing container environme

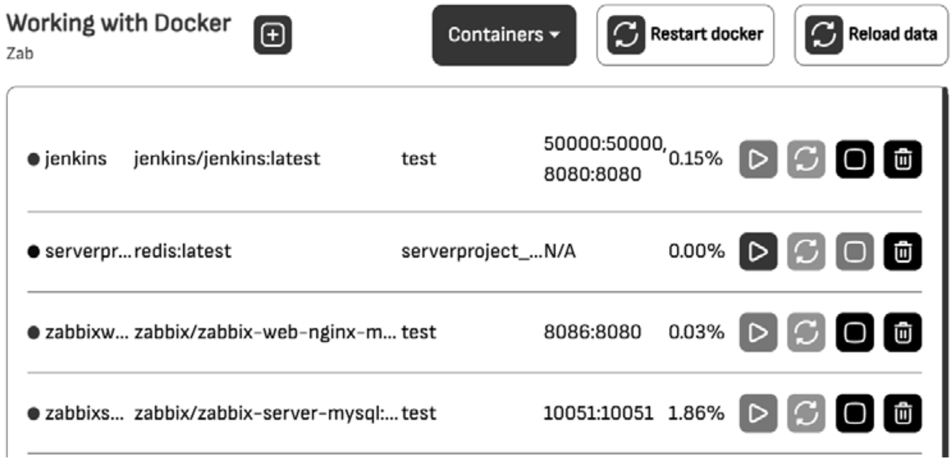


Fig. 8. Docker management page (containers, icons, network, volumes)

Manage system services and task scheduler through the web interface

Another salient feature is the capacity to oversee system services and the Task Scheduler via the web interface (Figs. 9, 10). This functionality enables users to view active services, initiate or halt their operation, and configure the task scheduler to automate the execution of tasks on servers. This solution facilitates the management of server operations, reduces the time required for configuration, and enhances productivity by centrally controlling all processes on disparate servers. Users can perform tasks such as adding, editing, and deleting scheduled tasks directly through interaction with the interface.

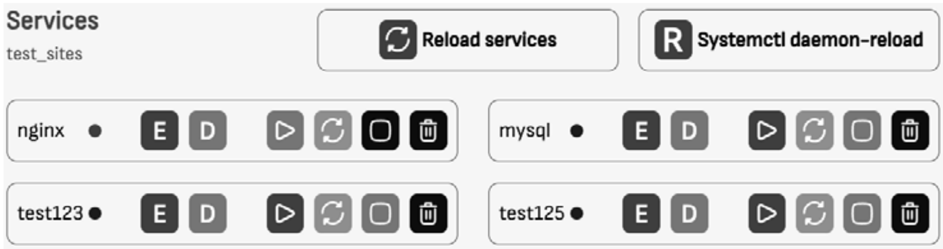


Fig. 9. Service management page

Advantages of Centralized Server Management

Centralized server management through a single web interface offers several significant benefits:

- *No Need for Constant Connectivity.* One of the primary advantages of using Paramiko through a web application is that it eliminates the need for a constant, stable connection to the server from the user side. Instead of a direct SSH connection from each device, administrators can work through a web interface located on a server in a region with reliable power supply and internet connection. This ensures that even in areas with intermittent internet connectivity, users can still send commands to the server.
- *Optimized Connectivity in Poor Internet Conditions.* Since the connection to the server is made via the centralized web application, users only need a short, stable connection to send commands. Even in poor internet conditions or frequent disconnections, the web application continues to process and send commands to the server, minimizing the need for constant user intervention. This is especially

beneficial in areas with unreliable connectivity, allowing administrators to execute critical tasks without the need for prolonged direct SSH connections.

- *Web Interface for Server Management.* Paramiko ensures SSH connectivity, while the web application serves as the interface between the user and the remote server. This allows administrators to remotely connect to the server's terminal through the web interface, ensuring stable connections and reducing reliance on the quality of the local network.
- *Faster Command Execution via the Web Application.* Sending commands through the web interface optimizes server operations even in low-speed internet environments. Since the web application is hosted in regions with better connectivity, commands are processed and dispatched more quickly. This reduces downtime and increases productivity by allowing administrators to quickly address various issues, even when local connections are unstable or limited.

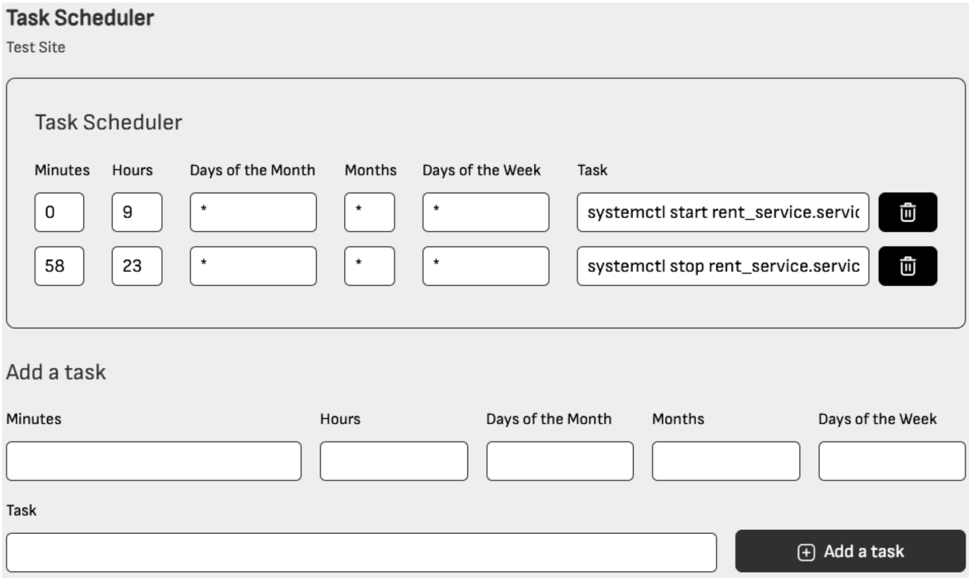


Fig. 10. Task Scheduler configuration

Administrators can manage the entire server infrastructure from a single place, which increases efficiency and reduces the need to constantly switch between different tools to manage various types of servers. This is especially useful in hybrid environments where both physical and cloud servers are used. Moreover, the centralized approach helps standardize administrative processes, lowering the risk of errors during routine tasks.

This approach also allows you to standardize administration processes, reducing the number of possible errors during routine operations. The interface provides an intuitive way to manage server configuration, task scheduler, services, updates, as well as server health monitoring and container management, which significantly reduces the level of technical knowledge required for infrastructure management.

Convenience of server process automation

The implemented system has effectively streamlined the automation of routine server processes through a unified web interface, thereby facilitating the management of heterogeneous servers. The system's integration with the APIs of the Hetzner cloud platform and AWS exemplifies its user-centric design, significantly simplifying the configuration and management of providers' infrastructure [16; 17]. The interface's comprehensive functionality enables administrators to execute all requisite actions without the need for dedicated applications or tools for each provider. The UI/UX design of the system prioritizes minimizing administrator actions, with all essential functions accessible within a few clicks in the web interface. For instance, instead of manually configuring the server through individual providers, the system enables the utilization of the "Ubuntu install Django Project" button, which employs a prescriptive configuration template, thereby reducing the time required for setting up the infrastructure by 44%.

The interface facilitates not only the utilization of predefined templates but also the creation of bespoke scripts or command sets for expedited setup. A segment of the recorded code for the command line interface (CLI) to install the requisite dependencies on a pristine server, along with an update and the installation of Python version 3.10.5 via pyenv, is presented below.

The addition of a command to the "General custom commands" group ensures its availability across all servers, irrespective of their location or type. However, for specific tasks, "Individual custom commands" have been developed, which are displayed exclusively within a single server (Figs. 11, 12).

Fig. 11. Adding a new general custom command

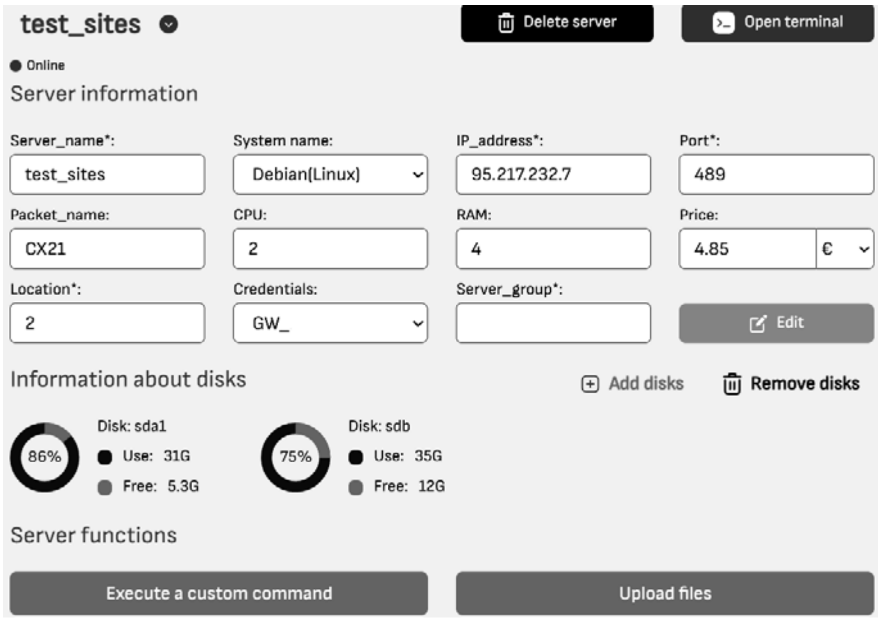


Fig. 12. Interaction page with the server

CASE STUDY

This section presents an analysis of the server infrastructure configuration process, comparing manual and automated methods. The purpose of this evaluation is to determine the effectiveness of automation in managing server environments. The study involved two types of servers: physical servers based on Raspberry Pi 4 and Hetzner CX21 cloud servers, both running Ubuntu 24.04. The time required to configure the infrastructure manually and using an automated system was measured and compared, highlighting the advantages and limitations of each approach. The evaluation focuses on the configuration processes, tools used, and the impact of automation on efficiency and accuracy.

Server Specifications

To evaluate the effectiveness of automating server infrastructure configuration, a case study was conducted using five Hetzner CX21 cloud servers and five physical servers based on Ubuntu 24.04 running on Raspberry Pi 4. Each physical server has its own IP address and SSH port for access. The time required for configuring the infrastructure manually and using the automated system was measured.

Characteristics of Physical Servers: Processor — Quad-core Cortex-A72 (ARM v8) 64-bit SoC @ 1.5GHz; RAM: 8 GB SDRAM; Disk: 32 GB; Operating System: Ubuntu 24.04.

Characteristics of Hetzner CX21 Cloud Servers: Processor -2 vCPU (virtual cores) AMD EPYC 7002 series; RAM: 4 GB RAM; Disk: 40 GB SSD; Operating System: Ubuntu 24.04.

Manual Configuration Process

In this case, the process involved running commands through the terminal without using Ansible or Terraform, as these tools require additional knowledge and expertise. Working with Ubuntu, the server contains only minimal pre-installed programs, which requires the engineer to manually execute the following steps:

- Connect to the server.
- Update the system.
- Install all necessary packages (web server, database server, etc.).
- Prepare directories and create a virtual environment.
- Copy the project.
- Configure the web server.
- Check the results.

A Bash script for Ubuntu servers was written in advance. After the cloud servers were created and the operating system was installed on the physical servers, the administrator needed to log into each server via SSH (without using third-party software, just the standard admin console and SSH tools), transfer the script to the server, set the appropriate permissions, and run it, hoping that all syntax and indentation were correct. Time was spent switching between the browser and terminal, looking for IP addresses, and more. As a result, the manual configuration of 10 servers took 18 minutes using the terminal and a script file.

Automation Configuration Process

For efficiency evaluation, the server configuration process was carried out using the developed web application. The web application includes a page for file editing and running scripts via an integrated code editor that supports a set of popular programming languages. This allows users to quickly edit configuration files and scripts, such as Terraform or Ansible, and run them directly from the browser on their servers. Syntax checking plays a crucial role in improving code-writing speed, as errors are immediately visible (Fig. 13).

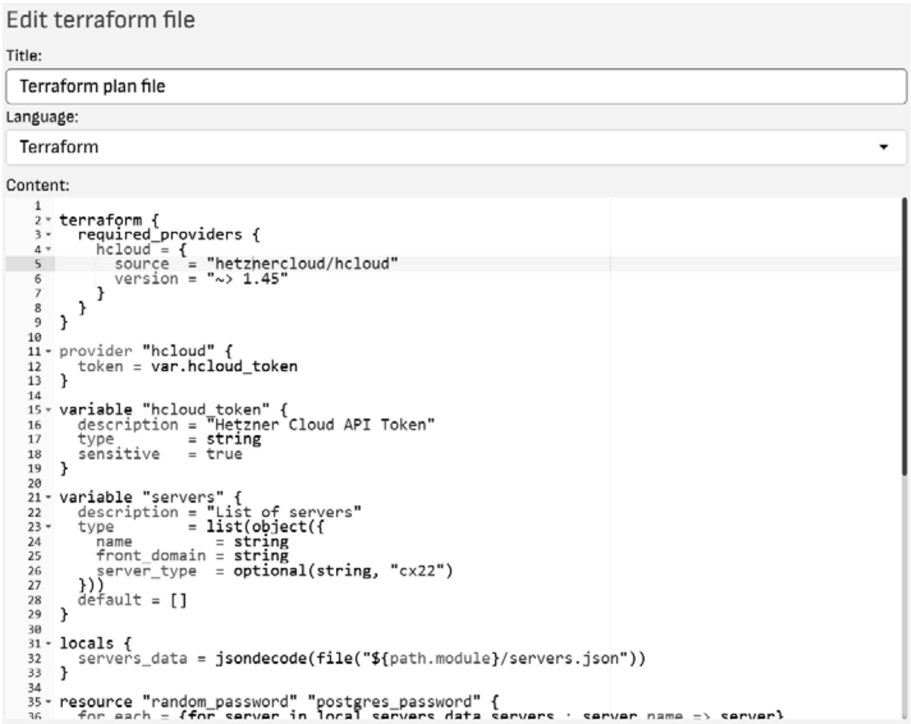


Fig. 13. Integrated file editor with syntax highlighting (programming languages)

For the time measurement in this case study, the configuration process was performed using the web application’s quick command tools without utilizing Ansible or Terraform. However, the web application already provides templates for working with these tools to facilitate interaction with them.

Once all the servers were created and the operating system was installed on the new physical servers, the following steps were performed within the web application:

- Add server authorization details.
- Add physical servers (cloud instances are displayed automatically via the cloud provider’s API).
- Open the first server (of any type) and use the “add custom commands” function to specify the script for execution (the newly created command will appear on all server pages, both existing and new).
- On each server page, click the newly created command and wait for the execution result.

As a result, the automatic configuration of 10 servers took 10 minutes using the web application and the script file. By using the web application, no additional time was spent on connecting, transferring script files, or entering commands manually.

To calculate how the automated system reduces the time required to configure server infrastructure, the following formula was used:

$$E_t = \frac{(T_m - T_a)}{T_m} * 100\% ,$$

where E_t — time savings percentage; T_m — manual configuration time (without automation); T_a — configuration time using the automated system.

The results obtained, measured using a timer, indicate that 18 minutes were spent manually configuring the servers. This time was obtained from the actual process of configuring the servers in a real environment, involving the following steps: manually connecting to each server, passing the configuration script, setting the appropriate permissions and executing the script, while frequently switching between the terminal and browser to look up IP addresses and other details. The total duration of the manual configuration of all ten servers was thus 18 minutes.Consequently, $T_m = 18$ minutes signifies the actual time taken to manually configure a server without automation. The 10 minutes required for the automated configuration were determined using the developed web application, with the automated system simplifying the process by executing the script on all servers, thereby greatly reducing the need for manual intervention. The total duration from the addition of servers to the execution of the customization script was 10 minutes. Therefore, $T_a = 10$ minutes signifies the time required to complete the server configuration using the automated system through the web application.

As a result, by arithmetic calculation we have $E_t = 44\%$.

Thus, the system reduces the configuration time by 44%, demonstrating the efficiency of automation:

- *Before automation:* 18 minutes for manual server configuration.
- *After automation:* 10 minutes using the web application (44% time savings).

The implementation of the automated solution, utilizing configuration templates, reduces the server configuration time by 44%.

DISCUSSION: RESULTS AND FUTURE RESEARCH

The study demonstrated that centralized management of both physical and virtual servers through a web interface significantly impacts the daily operations of administrators and DevOps specialists. Using such a system can considerably reduce the occurrence of errors during manual operations, as well as minimize unexpected server failures that require human intervention. It simplifies the management of complex infrastructures and reduces the time spent on routine tasks. Automation of these processes helps improve the overall efficiency of DevOps teams, enabling them to focus on critical aspects of software development and maintenance. Centralized solutions provide not only operational control over infrastructure but also promote rapid implementation of changes in response to dynamic task requirements, making this approach highly relevant to modern workflows and enhancing IT operations productivity.

Current solutions have certain limitations, most notably the lack of an integrated approach that allows centralized management of both physical and cloud-based servers through a single web interface. Although individual solutions allow partial integration of such servers, the absence of a unified platform capable of effectively managing both types of resources presents a significant drawback.

For instance, cloud providers offer built-in tools to manage only their services, but they do not provide solutions for managing physical servers or servers from other providers. This creates the need to use multiple tools to manage different infrastructures, complicating the process and reducing administrative efficiency. An integrated solution could streamline the management of hybrid infrastructures, reducing the time spent on administration and improving the overall reliability of the systems.

CONCLUSION

The study's development of a centralized hybrid infrastructure management system is notable for its provision of integrated administration of physical servers and cloud instances through a single web interface. A key feature of the system is its support for remote access via SSH using the Paramiko library, which facilitates convenient management of servers regardless of their location. The relevance of the proposed solution is evidenced by the contemporary requirements for server infrastructure, which necessitate a high degree of automation, convenient monitoring and management of a substantial number of cloud and physical resources.The presented system offers advanced administration capabilities, rendering it a versatile and powerful tool for DevOps professionals.

The study's findings indicate that the implementation of centralized solutions enhances the reliability of server infrastructure, streamlines maintenance operations, and mitigates the risk of human errors during manual management.

In the future, the centralized management system will be enhanced to include integration of additional cloud providers, such as Google Cloud Platform and DigitalOcean. This will extend administrative capabilities to a larger number of cloud platforms. The system may include improved monitoring features, such as tools for automatic notification of potential problems, load analysis tools, or detection of security vulnerabilities. It is expected to support automatic deployment of servers and containers using Kubernetes, which will provide greater scalability without increasing administrative costs. It is also of interest to optimize algorithms for more cost-effective and faster provisioning of servers in cloud providers using IaC (infrastructure as code) tools that support scalability.

Acknowledgments. The authors express their gratitude to the staff of the Institute of Computer Science and Artificial Intelligence at V.N. Karazin Kharkiv National University for their valuable discussions and insightful comments, which significantly contributed to improving the quality of the presented results.

REFERENCES

1. Hernandez L., C.E. Uc Rios, "Docker Optimization of an Automotive Sector Virtual Server Infrastructure," *Knowledge Engineering and Data Science*, 7(1), pp. 71–85, 2024. doi: https://doi.org/10.17977/um018v7i12024p71-85

2. X. Song, Y. Wu, Y. Yao, B. Li, "Research on Distribution Automation Testing Platform Based on Kubernetes Technology," *J. Phys.: Conf. Ser.*, 2777 012005, 2024. Available: https://iopscience.iop.org/article/10.1088/1742-6596/2777/1/012005/pdf
3. K. Nugroho, A.D. Abrariansyah, S. Ikhwan, "Perbandingan Kinerja Library Paramiko Dan Netmiko Dalam Proses Otomasi Jaringa," *InfoTekJar: Jurnal Nasional Informatika dan Teknologi Jaringan*, vol. 5, no. 1, 2020. Available: https://jurnal.uisu.ac.id/index.php/infotekjar/article/download/2758/pdf
4. P. Kaushik, A.M. Rao, D.P. Singh, S. Vashisht, S. Gupta, "Cloud Computing and Comparison based on Service and Performance between Amazon AWS, Microsoft Azure, and Google Cloud," *International Conference on Technological Advancements and Innovations* (*ICTAI*), pp. 268–273, 2021. doi: 10.1109/ ICTAI53825.2021.9673425
5. M. Spichkova, J. Bartlett, R. Howard, A. Seddon, X. Zhao, Y. Jiang, "SMI: Stack Management Interface," *23rd International Conference on Engineering of Complex Computer Systems (ICECCS)*, pp. 156–159, 2018. doi: 10.1109/ ICECCS2018.2018.00024
6. S. Bhatia, C. Gabhane, "Terraform: Infrastructure as Code," in *Reverse Engineering with Terraform*. Apress, Berkeley, CA, 2024, pp. 1–36. doi: https:// doi.org/10.1007/979-8-8688-0074-0_1
7. P. Kirkbride, "Using SSH," in *Basic Linux Terminal Tips and Tricks*. Apress, Berkeley, CA, 2020. https://doi.org/10.1007/978-1-4842-6035-7_5
8. A. Witt, S.Westling, *Ansible in Different Cloud Environments*, 2023. Available: http://www.diva-portal.org/smash/get/diva2:1765141/FULLTEXT01.pdf
9. A. Chan, D. Macpherson, *Red Hat Satellite 6 System Architecture*. Accessed on: 05.01.2024. [Online]. Available: https://docs.redhat.com/en/documentation/red_hat_ satellite/6.0/html-single/user_guide/index#sect-Red_Hat_Satellite-User_Guide-Red_Hat_Satellite_6_System_Architecture
10. K.I. Mantilla, A. Flórez, "Development of a Monitoring Module for Physical and Virtual Servers: Advanced Computing Center Case of the Universidad Pontificia Bolivariana, Bucaramanga, Colombia," *Journal of Physics: Conference Series*, vol, 1513, 2020. Available: https://iopscience.iop.org/article/10.1088/1742-6596/1513/ 1/012006
11. Almeida Wilfred, "SC Connect: Secure Server Access from Mobile Device," *IEEE Bombay Section Signature Conference (IBSSC)*, pp. 1–3, 2022. doi: 10.1109/IBSSC56953.2022.10037430
12. B. Donald, D. Quirk, "Managing Infrastructure: DCIM: Is it a Product or a Solution?" *ASHRAE Journal*, vol. 57, pp. 82–84, 2015. Available: https://www.researchgate.net/publication/285366410_DGIM_Is_it_a_Product_or_a_ Solution_Managing_Infrastructure
13. M. Zadka, "Paramiko," in *DevOps in Python*. Apress, Berkeley, CA, 2019. doi: https://doi.org/10.1007/978-1-4842-4433-3_9
14. A. Florez, L. Serrano, U. Gomez, L. Suarez, H. Villalarga, H. Rodriguez, "Analysis of Dynamic Complexity of the Cyber Security Ecosystem of Colombia," *Future Internet*, 8(3), 33, 2016. doi: https://doi.org/10.3390/fi8030033
15. K. Maurice et al., "Hello from the Other Side: SSH over Robust Cache Covert Channels in the Cloud," *NDSS*, vol. 17, pp. 8–11, 2017. doi: http://dx.doi.org/10.14722/ndss.2017.23294
16. C. Tao, G. Xiaofeng, C. Guihai, "The Features, Hardware, and Architectures of Data Center Networks: A Survey," *Journal of Parallel and Distributed Computing*, 96, 45, 2016. doi: 10.1016/j.jpdc.2016.05.009
17. Y. Su, T. Zhou, W. Zheng, L. Zhao, H. Zhen, G. Huang, "Construction of Power System Computing and Analysis Platform Based on Cloud Computing," *Southern Power System Technology*, 16(07), pp. 67–75, 2022.

Received 27.01.2025

INFORMATION ON THE ARTICLE

Ivan S. Byzov, ORCID: 0009-0004-2950-7814, V.N. Karazin Kharkiv National University, Ukraine, e-mail: utelephona@gmail.com

Sergiy V. Yakovlev, ORCID: 0000-0003-1707-843X, V.N. Karazin Kharkiv National University, Ukraine, e-mail: s.yakovlev@karazin.ua

СПРОЩЕНЕ КЕРУВАННЯ ФІЗИЧНОЮ ТА ХМАРНОЮ ІНФРАСТРУКТУРОЮ ЧЕРЕЗ ЦЕНТРАЛІЗОВАНИЙ ВЕБ-ІНТЕРФЕЙС / І.С. Бизов, С.В. Яковлев

Анотація. Присвячено розробленню системи централізованого управління фізичними та віртуальними серверами через веб-інтерфейс. Система дозволяє адмініструвати серверну інфраструктуру за допомогою різних інструментів. Означено ключові компоненти системи, інтеграція з інструментами управління серверами, використання Paramiko для віддаленого доступу через SSH, а також можливості застосування інших технологій для управління хмарними і фізичними серверами через веб-інтерфейс. Система дає змогу централізовано керувати різними типами серверів через веб-інтерфейс, підтримуючи як базові, так і розширені адміністративні операції. Інтеграція з Docker і підтримання хмарних API забезпечують зручну автоматизацію і спрощують операції із хмарними і фізичними серверами. Система слугує універсальним інструментом для DevOps-інженерів. Дослідження підкреслює актуальність і важливість централізованого підходу до управління серверною інфраструктурою, що підвищує ефективність і надійність роботи серверів у сучасних умовах.

Ключові слова: централізоване управління, SSH, веб-інтерфейс, хмарні сервіси, автоматизація, DevOps.

UDC 004.855.5
DOI: 10.20535/SRIT.2308-8893.2025.2.03

NAVIGATING CHALLENGES IN DEEP LEARNING FOR SKIN CANCER DETECTION

V. NIKITIN, V. DANILOV

Abstract. Skin cancer is one of the most prevalent malignancies worldwide. A critical factor in reducing mortality rates is the early detection. It underscores the need for accessible Computer-Aided Diagnostic (CAD) systems. Recent advancements in Deep Learning (DL) have shown great promise in addressing this challenge. Despite this progress in the field of machine learning, researchers encounter numerous obstacles when it comes to skin cancer classification. This article examines the current state of DL-based skin cancer diagnostics. Critical aspects of system development, including data preprocessing, model training, and performance evaluation, are addressed. Moreover, the article highlights opportunities for innovation that could significantly advance the field. By providing a comprehensive overview, this article aims to guide researchers and practitioners in optimizing DL models, addressing existing limitations, and exploring emerging trends to enhance diagnostic accuracy and accessibility.

Keywords: skin cancer, deep learning, classification, transformers, CNN, GAN, data preprocessing, data augmentation.

INTRODUCTION

Skin cancer affects millions of individuals every year. Malignancy originates in the skin's primary layers — the epidermis, dermis, and hypodermis. Due to its exposure to ultraviolet (UV) radiation, the epidermis is the most common site for skin cancers. Keratinocytes and melanocytes in these layers can undergo mutations due to prolonged exposure to UV radiation, which can lead to malignancy. The three major types of skin cancer include Basal Cell Carcinoma (BCC), Squamous Cell Carcinoma (SCC), and Malignant Melanoma [1].

BCC is the most common skin cancer, typically presenting as a slow-growing, pearly lesion. While it rarely metastasizes, untreated BCC can cause significant local tissue damage. SCC, the other type, is more aggressive and capable of metastasis. It often appears as scaly, crusty lesions, especially on sun-exposed skin, SCC requires timely intervention to prevent systemic spread. Malignant melanoma, though rare, is the most aggressive form of skin cancer and is prone to rapid metastasis [1].

Regardless of the type of skin cancer, early detection is crucial for successful treatment. Still, there are many barriers preventing timely diagnosis. According to the World Health Organization, there is a global shortage of dermatologists, particularly in low- and middle-income countries [2]. Studies show evidence that lower-income populations are more likely to have advanced stages of skin cancer [3]. Hence, obtaining professional help to determine whether a lesion is malignant can be challenging. Even in developed regions, busy lifestyle and/or financial constraints can prevent individuals from seeking medical attention.

Computer Aided Diagnostics (CAD) system is a solution for assisted or automated diagnostics. They can be especially helpful in reducing human errors and expanding access to advanced medical care. Deep Learning (DL) with its advancements in recent years is a promising solution for skin cancer CAD. Machine Learning (ML) models can help to detect malignant lesions early via specialized or common hardware, potentially covering gap in dermatological services. Such systems can be particularly useful as prescreening tools, indicating cases where professional help is necessary.

DL has revolutionized various areas over the recent years, including medical imaging. During the past decade, researchers have developed numerous DL models for skin cancer classification with varying levels of success. This diversity of models and methods creates a sophisticated landscape to navigate. The article aims to underline key challenges and obstacles that researchers face developing skin cancer DL model as of 2024, providing examples of solutions proposed in existing researcher. Finally, it highlights areas for improvement that could take current DL-based CAD systems to a new level.

To simplify navigation, the article is divided into sections, each focusing on a critical component in the development of DL-based CAD systems for skin cancer detection.

1. *Researchers and Communities*: An overview of labs, teams, and organizations that have made significant advancements in DL-based skin cancer tasks in recent years.

2. *Datasets and Data Challenges*: A list of widely used datasets with their descriptions. This section also includes an overview of data based obstacles.

3. *Development Pipeline*: Addressing data normalization, augmentation, and techniques to fix data challenges highlighted in the previous section.

4. *Model Training and Optimization*: This section explores the complexities of selecting model architectures, leveraging transfer learning, and optimizing hyperparameters.

5. *Evaluation and Validation*: Various metrics suitable for the task are explored in detail.

6. *Opportunities for Innovation*: Directions and areas that could significantly enhance the quality of DL-based CAD systems.

7. *Guidance for Emerging Researchers*: Practical advice for newcomers to the field, including best practices in research design, navigating the literature, and identifying emerging trends.

By exploring these themes, the article provides a concise resource highlighting the current challenges, solutions, and future directions in the application of DL to skin cancer detection. Ultimately, the aim is to contribute to the development of more accurate, accessible diagnostic tools that can improve patient outcomes and reduce healthcare disparities.

RESEARCHERS AND COMMUNITIES

Stanford University Team. In 2017, the study "Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks" was published in Nature by a team from Stanford University. They researched a dataset of approximately 130,000 clinical images representing over 2,000 different skin diseases. They took Google Inception v3 CNN architecture, pre-trained on the ImageNet dataset, and fine-tuned it for skin lesion classification. Performance was validated against a group of 21 dermatologists, ensuring clinical relevance. It turned out the model achieved performance on par with board-certified dermatologists [4].

International Skin Imaging Collaboration. The International Skin Imaging Collaboration (ISIC) is the largest consortium that made a significant contribution to the skin cancer classification problem. The collaboration created the largest publicly available skin cancer dataset and conducted multiple challenges to encourage the development of advanced algorithms for skin lesion analysis. In mid-2024, ISIC released its latest dataset [5]. ISIC promotes transparency and reproducibility in dermatological research while safeguarding patient privacy. Its extensive use in numerous studies has made it a cornerstone for developing and validating DL models for skin cancer diagnostics, significantly accelerating progress in the field.

Memorial Sloan Kettering Cancer Center. Memorial Sloan Kettering Cancer Center (MSKCC) is one of the world's leading institutions dedicated to cancer treatment, research, and education. In dermatology, MSKCC implemented AI to support the detection of melanoma. The institution is actively working on implementing CAD systems to improve the effectiveness of diagnostics. They consider AI a complementary tool for medical experts [6].

Google Health and DeepMind. Google's research teams actively worked to apply DL to medical images. For instance, Yuan Liu and Peggy Bui published a paper titled "A Deep Learning System for Differential Diagnosis of Skin Diseases" in 2019 [7]. They developed a system that was able to accurately differentiate 26 common skin conditions.

Individual Researchers. There are numerous ML enthusiasts and experts working on the problem all over the world. For example, Yilmaz et al. achieved an accuracy of 82% with on the ISIC 2017 dataset with the NASNetMobile model [8]. Baygin et al. integrated textural and deep features to develop a pyramidal hybrid model receiving a classification accuracy of 91.54% with 10-fold cross-validation [9]. Agarwal and Singh were able to get 86.65% accuracy on ISIC Archive utilising convolutional neural networks with transfer learning [10].

Many more studies report remarkable results with the models and the data used. Each year, hundreds of new articles on the topic are being published, some of them making breakthroughs through usage of new technology or metric. The article mentions key pioneering research as examples of successful problem-solving for medical imaging. The main idea of this article is to guide through the current landscape of skin cancer images with DL leading to even more advances in the field.

DATASETS AND DATA CHALLENGES

Data is the base of any ML algorithm. The success of any DL project heavily depends on access to robust and comprehensive datasets as well as the correct usage of those. By processing the data, the model extracts the most important features from it so that it can make correct predictions on unseen data of a similar nature. D. Wen et al. (2022) made an exhaustive list of datasets that were publicly available at the time [11]. The section below mentions key datasets from that list, adding the data that was released since 2022.

Datasets. *ISIC Archive.* The ISIC Archive [5] is the largest data repository with cancerous skin images so far. As of today, the number of open-for-usage samples is nearing half a million. More than 400,000 of those are images released as part of the latest ISIC Challenge — ISIC 2024 [12]. The archive includes a diverse range of skin lesions: melanomas, nevi, BCCs, SCC, etc. It includes both annotated and unannotated data sourced from various international centers. Some of the images contain precise borders of the lesion, allowing not only classification but segmentation as well. It might be the most impactful contribution to DL-based skin cancer CAD systems in terms of data.

HAM10000. The dataset [13], also known as "Human Against Machine with 10,000 training images" (HAM10000), published by Philipp Tschandl et al. (2018), is the most popular dataset for DL skin cancer classification. It contains a total of 10,015 dermoscopic images, categorized into seven common types of skin lesions: actinic keratoses, basal cell carcinoma, benign keratosis-like lesions, dermatofibroma, melanoma, melanocytic nevi, and vascular lesions. Each image is labeled with a confirmed diagnosis, verified through histopathology, follow-up examinations, or expert consensus. It also includes metadata detailing patient demographics such as age, gender, and lesion location, which adds context for model training.

Due to its high-quality data, HAM10000 was used in numerous articles, normally combined with transfer learning to compensate for the lack of size. For instance, A.T. Priyeshkumar et al. (2024) developed an ensemble DL model for skin lesion classification using HAM10000, achieving high accuracy in differentiating between lesion types [14]. T.M. Alam et al. (2022) proposed a novel CNN architecture trained on HAM10000, focusing on improving classification performance through data augmentation techniques [15]. H.A. Owida et al. (2024) investigated the use of transfer learning with pre-trained models on HAM10000 to enhance melanoma detection [16]. A. Ameri et al. (2020) explored feature extraction methods using HAM10000 to improve the interpretability of DL models [17]. S.S. Chaturvedi et al. (2020) conducted a comparative study of DL algorithms on the HAM10000 dataset to identify the most effective approaches for skin lesion classification [18].

Non-ISIC Data. While the ISIC Archive and HAM10000 datasets are widely used in recent skin lesion research, several other datasets have historically contributed to the field, particularly before the availability of ISIC data.

The **PH²** dataset [19], developed by Pedro Hispano Hospital in Portugal, contains 200 dermoscopic images focusing on melanocytic nevi and melanomas. Despite its small size, the dataset is highly valued for its detailed manual segmentation masks and clinical annotations that include colors and dermoscopic structures. It is often used in studies emphasizing precise lesion segmentation and the analysis of dermoscopic features.

The **MED-NODE** dataset [20], created by the University of Twente in the Netherlands, includes 1,700 clinical (non-dermoscopic) images of melanomas and benign melanocytic nevi. Diagnoses in this dataset are confirmed through histopathology. It is particularly useful for research focusing on classification based on clinical photographs rather than dermoscopic images. This dataset presents challenges like variations in lighting, skin tone, and image quality, making it a valuable resource for developing models that can handle real-world conditions.

The **Dermofit** Image Library [21], developed by the University of Edinburgh, includes approximately 1,300 high-resolution images spanning 10 skin lesion categories. However, it is accessible only under a commercial license, which can limit its availability for some researchers. The library is often used in studies requiring high-resolution images and advanced feature extraction techniques.

The **SD-198** dataset [22], compiled from dermatology atlases, contains 6,584 images across 198 different skin disease classes. It is particularly valuable for multi-class classification research and for developing models that can identify rare skin conditions.

These datasets, while smaller in scale compared to ISIC, offer unique advantages by addressing diverse imaging conditions and focusing on specific research challenges, continuing to play a crucial role in dermatology research.

Data Challenges. *Class Imbalance.* Datasets often have a disproportionately large number of benign lesions compared to malignant ones. Such skewed data causes model bias and poor sensitivity in detecting rare but critical cases like melanomas. This problem is typical for skin cancer data. For instance, in the ISIC 2020 dataset, the benign nevi-to-melanoma ratio is approximately 9 to 1 [5]. If not handled properly, this imbalance can lead to high overall accuracy but poor performance when detecting minority classes, which are crucial for early diagnosis.

Underrepresentation of Diverse Skin Types. In recent years, it was pointed out that available skin cancer data is insufficient in terms of darker skin tone representation. It can affect a model's accuracy across different ethnicities and exacerbate health disparities [11], [23]. The research community considers this a serious problem as, even though skin cancer is less prevalent among individuals with darker skin, it is often detected at later stages, leading to drastic outcomes. For instance, Black individuals are three times more likely to be diagnosed with late-stage melanoma [24]. One of the reasons is that melanoma for that part of the population sometimes appears on less visible areas, such as the palms, soles, or under the nails, which are less frequently examined [25].

Privacy Concerns and Data Sharing Limitations. Skin cancer imaging data is part of a person's personal information. Consequently, it falls under many laws and acts aimed at protecting users' data. As this data is of medical nature, it complicates the situation even more. Among the commonly known regulations are:

- Health Insurance Portability and Accountability Act [26] (HIPAA) in the United States.
- The General Data Protection Regulation [26] (GDPR) in the European Union.

These regulations impose strict guidelines on the handling, sharing, and processing of medical data. Understanding these restrictions is crucial for researchers aiming to access and utilize dermatological datasets while ensuring compliance with legal and ethical standards.

Reluctance to Share Proprietary Datasets. Self-collected datasets offer researchers distinct competitive advantages. This method of data acquisition allows influencing the data flow from the very start, forming a more precise understanding of the data nature. At the same time, exclusive access to this data lets institutions uniquely fine-tune algorithms for enhanced performance, potentially leading to breakthroughs in skin cancer detection and diagnosis. Sulaiman Khan et al. (2022) [28] found that researchers often use private datasets exclusively or combine them with open data to achieve superior results in skin cancer classification tasks. Other reasons to withhold data might range from ethical to licensing concerns. Another issue is data privacy. This practice highlights the advantages of proprietary data but also underscores a significant barrier to progress in the field. When models are trained on private datasets, it becomes challenging to replicate studies, validate findings, or compare the effectiveness of different methodologies.

Data Labeling. Many of the publicly available skin cancer image samples are currently unlabeled, which makes them impossible to use with supervised DL algorithms. For instance, the ISIC 2020 dataset [5] includes 10,982 images without assigned classes. Meanwhile, accurately annotating medical images requires a high level of domain expertise, making the process both time-consuming and costly.

Artefacts in Image Acquisition. The most common way in which skin cancer images are obtained is dermatoscopy. It is a non-invasive imaging technique that involves examining and capturing skin lesions via a dermatoscope. Nevertheless, images of a dataset are often collected in different centers and institutions that utilize instruments with varying characteristics. However, when all the data is collected by a single facility, it may cause variations in resulting images, leading to models overfitting to device-specific artifacts instead of focusing on generic lesion features. Studies have shown that models trained on homogeneous datasets perform poorly on images from diverse sources [29], [30].

Ethical Considerations. Besides norms and regulations, one should realize that medical data is deeply personal. Processing it unavoidably raises some ethical concerns. Syed F.H. Shah in “Ethical implications of artificial intelligence in skin cancer diagnostics: use-case analyses” (2024) found that existing skin cancer analysis solutions require a great deal of transparency and collaboration to avoid potential ethical problems and misuse [31].

To resolve the ethical issues, the research community must show awareness when handling such sensitive data. In the paper “Ethical considerations for artificial intelligence in dermatology: a scoping review,” Emily R. Gordon (2024) identified key principles to follow to avoid ethics-related risks like fairness, inclusivity, transparency, accountability, etc. [32].

Additional Clinical Criteria. While lesion images serve as the primary source of information, various studies suggest that including more contextual data can positively impact classification performance. Esteva et al. (2017) noted that including clinical data could further enhance performance [4]. Haenssle et al. (2018) demonstrated that combining dermatoscopic images with patient information improved melanoma detection rates [33]. Pacheco and Krohling (2019) found that integrating clinical metadata with images in DL models led to higher accuracy in skin lesion classification [34]. These additional criteria may vary: patient age is important because certain age groups are more inclined to have some types of skin cancer than others, evolution over time is one of the key characteristics that experts use to make a decision on lesion state, as well as the size of the lesion. Embedding mathematical measures derived from the image, such as fractal dimension, may also contain insights useful for correct classification [35].

Incorporating these features into a classification model is a new challenge, as image and contextual data normally have different modalities. Making a model more complex to handle this may cause unintended overfitting. Therefore, the challenge is to balance model complexity while integrating features effectively.

Preprocessing Stage. In the previous section the key data challenges were identified. Many of them such as class imbalance, ethnic groups underrepresentation, restricted data, unlabeled data, inconsistency in the data gathering tools and noise can be addressed on a data preprocessing stage. In this section the key steps of data preparation are discussed.

Balancing Data. Some datasets do not suffer from class imbalance issues. For example, Kaggle dataset for malignant melanoma classification [36] contains approximately 5000 samples of each class which makes it a good option for fine tuning lightweight melanoma classifiers for exploration and education purposes.

Nevertheless, most of the available data samples are severely imbalanced with the largest amount of lesions being benign. For instance, HAM1000 dataset contains around 60% of benign nevus samples. Provided that this dataset is used for multi class classification, it makes it very imbalanced

Data Augmentation. A typical method for handling imbalanced data involves performing augmentations on the training data. Shen (2022) proposes an effective way of augmenting data for DL skin cancer classification with a significantly reduced search space of 60 possible transformations, compared to existing methods like AutoAugment and RandAugment [37]. Himel et al. (2024) applied data augmentation to 3,295 malignant images in the HAM10000 dataset, increasing the number to 5,000 [38]. These augmentations included rotation, flipping, and zooming. However, their approach may seem somewhat straightforward, as they simultaneously disregarded more than 1,500 benign lesion images to balance the data. This was also risky, as it essentially balanced the data through data loss. Polat et al. (2020) augmented images of the same dataset with noise, scaling, and rotation, increasing the total number of images to more than 33,000 [39]. Walker et al. (2019) used cropping in addition to the aforementioned methods while working with the ISIC 2017 Challenge [5], [40]. Ali et al. (2021) applied color-shifting using principal component analysis to create augmented images [41].

Another way of augmenting images is application of Generative adversarial network (GAN) architecture. Wu et al. (2020) did a review of GAN application for augmenting skin cancer images to solve data imbalance problem [42]. In article they note that various GAN architectures, including DCGANs [43], style-based GANs [44], TED-GAN [45], SPGGAN [46], and conditional GANs (CGANs) [47], have been employed to generate high-quality, diverse synthetic images. Enhancements such as artifact removal, attention modules, informative noise vectors, and stability improvements like the Two-Timescale Update Rule have further optimized GAN performance. Wu state that these methods have consistently improved classification metrics — including accuracy, sensitivity, specificity, and F1 scores — by providing richer training data and more reliable image generation, ultimately enhancing the effectiveness of skin cancer classification models.

While this approach offers benefits, it also comes with potential risks and drawbacks. Excessive noise, extreme scaling, or unusual rotation angles can distort images to the point where essential features are obscured or altered. If the augmentation methods produce images that are too similar to each other, the model might overfit to these synthetic variations rather than learning generalized features. This reduces the model's ability to perform well on truly unseen data. Adding noise can sometimes degrade image quality to a level where the model's ability to extract meaningful features is compromised. Drawbacks also include increased computational cost of creating and processing larger amount of data.

Fixing Ethnic Representation Issues. To mitigate the issue of underrepresentation, several strategies can be employed to improve both the diversity of data and the fairness of DL models.

- Color jittering, which adjusts the brightness, contrast, and saturation of images, can simulate a broader range of skin tones.
- Histogram equalization can improve the visibility of features across different skin tones.
- Style transfer can modify images to appear as if they belong to varied skin types.

Pope et al. (2024) compared training model on imbalanced and sampled datasets [48]. However, their results show that, albeit the sampled model tends to be less biased towards dark-toned skin, the overall accuracy decreases.

Beyond augmentation, synthetic data generation can further expand datasets, particularly for underrepresented skin tones. GANs are useful for generating synthetic images that mimic lesions on darker skin, while domain adaptation techniques can align feature representations to ensure models perform well across different skin tones. Rezk et al. (2022) composed a skin cancer dataset from existing open access data and applied style transferring to make it more diverse in terms of skin tone [49].

There also were instances of less traditional approach to solve this problem. Continuing their study of 2019, Walker et al. (2024) used sonification of skin cancer data to decrease influence of skin tone on classification [50]. The analysis demonstrated high and comparable diagnostic accuracy for both fair skin (FS) and darker skin (DS), with ROC curve AUCs of 0.858 and 0.856, respectively, indicating no significant difference between the two groups. Sensitivity and specificity values were similar (around 80–85%), confirming the model's consistent performance in detecting true positives and negatives across diverse skin tones. The results suggest that the classifier maintains equivalent diagnostic reliability for both FS and DS, even with the limitations of smartphone-based imaging.

To ensure fairness in model performance, algorithmic adjustments can be utilized. Reweighting samples can help balance training by assigning higher weights to underrepresented classes, while bias correction layers can be integrated into models to correct inherent biases.

Data Cleaning. Data cleaning is a crucial step in DL, involving the preparation and preprocessing of data to improve its quality for model training [51]. This process includes various data processing techniques that enable more effective feature extraction, ultimately enhancing model performance.

Normalization. To address inconsistencies in the conditions under which images of skin lesions are captured, data normalization and denoising are essential. Normalization techniques, such as Gray World, Shades of Gray, or max-RGB algorithms, adjust the color balance of images to compensate for lighting differences [41], [52]. Global or adaptive histogram equalization methods, like Contrast Limited Adaptive Histogram Equalization (CLAHE), can be applied to improve contrast and standardize the intensity distribution across images [53], [54]. Gamma correction is another option, adjusting non-linear luminance or color values to standardize brightness and contrast.

Standard normalization methods, such as scaling pixel values to have zero mean and unit variance (z-score normalization) or scaling between 0 and 1 (min-max normalization), are commonly used in image preprocessing [55]. For example, 0–1 normalization was utilized in [41]. These steps ensure that images fed into DL models are consistent and of high quality.

Noise Reduction. Removing noise from images is critical to enhance quality and improve model accuracy. Common strategies include applying filters like median or Gaussian filters to smooth images and reduce noise. Midasala et al. (2024) applied a top-hat transform to remove thick hairs, while filters effectively eliminated noise and thin hairs, albeit with limitations on contrast-based histogram equalization [56]. Morphological operations, such as opening and closing, assist in removing small artifacts.

Body hair presents a specific challenge in analyzing skin images, as it can obscure important lesion features. Algorithms like the DullRazor detect hair pixels using edge detection and inpaint these regions to remove hair from the image [57]. Thresholding methods identify hair regions so that they can be cleaned out.

Another approach to denoising is the use of autoencoders. These models learn to create a compressed representation of the input and reconstruct it, effectively removing noise in the process. Maurya et al. (2024) used autoencoders for denoizing and feature extraction [58].

Segmentation. Segmentation is another crucial preprocessing step, involving the selection of the region of interest (ROI) from the image—in this case, the skin lesion. Accurate segmentation focuses analysis on the lesion and removes irrelevant background information. This topic is well-researched, and researchers often utilize large pretrained models.

For instance, Himel et al. (2024) [38] used Meta's Segment Anything Model [59] to perform segmentation on skin cancer images from the HAM10000 dataset. They then used the segmented images to pass only the ROI to feature extraction and classification models, achieving 96% accuracy using Google's Vision Transformer (ViT) [60]. This approach demonstrates the effectiveness of combining advanced segmentation models with powerful classification architectures.

TRAINING AND OPTIMIZATION

Model training and optimization are pivotal in developing robust DL models for skin cancer classification. These processes involve selecting suitable architectures, optimizing learning algorithms, fine-tuning hyperparameters, and employing strategies to enhance model performance while addressing overfitting.

Transfer Learning. Transfer learning is extensively utilized in medical image analysis due to the scarcity of labeled datasets [11]. By leveraging models pre-trained on large-scale datasets like ImageNet, researchers can fine-tune these models for specific tasks such as skin cancer classification [4]. This approach not only accelerates convergence but also often yields superior performance compared to training models from scratch.

In skin cancer classification, the data available is limited for many reasons. Therefore, usage of pretrained models is popular approach. Ali et al. (2021) [41], while developing a custom NN for skin cancer classification, applied transfer learning to the task with pre trained ResNet [61], AlexNet [62], VGG-16 [63], DenseNet [64] and MobileNet [65] achieving top accuracy of 86.09. Another example is using ViT in [60] with 0.96 accuracy result.

Optimization Techniques. *Learning Rate Scheduling.* Adjusting the learning rate during training can significantly influence model convergence. Techniques such as Step Decay, Exponential Decay, and more advanced methods like Cyclical Learning Rates and Warm Restarts are employed to optimize training efficiency [66, 67].

Optimizers. Selecting an appropriate optimization algorithm is crucial for training deep neural networks. Optimizers like Stochastic Gradient Descent (SGD) with momentum, Adam, and RMSProp are widely used. Adaptive optimizers like Adam combine the advantages of AdaGrad and RMSProp, providing efficient training for deep models [68]. In skin cancer classification tasks, Adam is often preferred for its ability to handle sparse gradients and noisy data (for instance in [69]).

Regularization Methods. Regularization techniques prevent overfitting by adding constraints to the model. L1 and L2 regularization add penalties to the loss function based on the magnitude of weights. Dropout randomly deactivates neurons during training, reducing interdependent learning among neurons. Batch normalization normalizes layer inputs, accelerating training and improving generalization.

It is also worth mentioning that on this stage is possible to address class imbalance problem if it was not solved completely on data preprocessing stage. Le et al. (2020) used focal loss to train on imbalanced data of HAM10000 and received 0.94 top accuracy with pretrained EfficientNetB1 [70].

Hyperparameter Tuning. Hyperparameters such as learning rate, batch size, network depth, and activation functions significantly impact model performance. Systematic methods like Grid Search and Random Search explore combinations of hyperparameters, while Bayesian Optimization offers a more efficient search by modeling the performance as a probabilistic function.

Same time hyperparameter tuning can always be costly no matter the technique chosen. In [69] it was done on only 10% of training dataset which allowed to select appropriate values more efficiently.

Early Stopping. Early stopping halts training when the validation loss stops decreasing, preventing the model from overfitting to the training data. This method is especially useful when training deep networks on limited datasets, as is common in medical imaging.

Ensemble Learning. Ensembling combines predictions from multiple models to improve robustness and accuracy. Techniques like averaging, majority voting, or stacking can enhance performance in skin cancer classification. Ghosh et al (2024) [71] utilized ensemble learning with DCNN [72], Caps-Net [73] and ViT [60].

Non Supervised Learning. *Semi-supervised learning*. Semi-supervised learning combines a small amount of labeled data with a large amount of unlabeled data during training. This approach is particularly beneficial in medical imaging, where labeling is expensive and time-consuming. Liu et al. (2020) employed a semi-supervised learning approach using a mean teacher model to leverage unlabeled skin lesion images, improving classification performance [7]. Techniques:

- **Pseudo-Labeling**: Assigning labels to unlabeled data using the model's own predictions and then retraining the model with this expanded dataset. This iterative process can improve performance but may propagate errors if the initial model is not accurate.

- **Consistency Regularization**: Encouraging the model to produce similar outputs when input images are perturbed or augmented, leveraging unlabeled data to learn robust features.
- **Mean Teacher Model**: Utilizing a student-teacher framework where the teacher model generates targets for the student model, which learns from both labeled and unlabeled data. A noisy stutent algorithm was successfully used in ISIC Kaggle competition in melanoma classification challenge [74].
- **Graph-Based Methods**: Modeling data as graphs where nodes represent samples and edges represent similarities, propagating labels through the graph to infer labels for unlabeled data.

Self-Supervised Learning. Self-supervised learning is a form of unsupervised learning where the model learns representations by solving pretext tasks created from unlabeled data. This approach is gaining traction in medical imaging. Chaitanya et al. (2020) showed that self-supervised learning on unlabeled medical images significantly improves model performance on downstream tasks with limited labeled data [75]. Techniques:

- **Contrastive Learning**: Learning representations by maximizing agreement between differently augmented views of the same data sample. The key idea is to develop a non-trivial network that preserves similar semantic structure for two (somewhat modified) versions of the same image and keeps two different images distinguishable [76]. Models like SimCLR have been adapted for medical images to learn robust features from unlabeled data [77].
- **Pretext Tasks**: Designing tasks such as predicting image rotations, solving jigsaw puzzles, or reconstructing distorted images to force the model to learn meaningful features. Haggerty and Chandra (2024) showed that models pre-trained using SSL (Barlow Twins) significantly outperformed those pre-trained with SL on ImageNet in scenarios with limited labeled data specifically using skin cancer images. Moreover, by applying additional SSL pre-training on smaller, task-specific datasets (like skin lesion images), SL-pre-trained models could achieve performance equivalent to SSL models. This demonstrates that even minimal further SSL pre-training can be as effective as extensive pre-training on large datasets [78].

EVALUATION AND VALIDATION

Next important step in the DL pipeline is assessing the model's effectiveness. It involves measuring the model's performance on previously unseen data. Testing dataset is distinct from the validation dataset used during training. The validation dataset helps evaluate the model's progress and detect overfitting, enabling techniques like early stopping to be applied.

To accurately identify model quality, an appropriate evaluation strategy and metrics must be chosen. For instance, when working with relatively small training and testing datasets, it's crucial to focus on validating whether observed improvements stem from the new approach rather than statistical noise.

Selecting suitable metrics is equally important. In medical diagnostics, for example, both type I and type II errors carry significant consequences. However, it is generally considered "better" to classify a healthy patient as sick than to miss diagnosing an ill patient. Moreover, medical datasets, such as those related to skin cancer, are often highly imbalanced. In such cases, relying solely on accuracy as the primary metric can be misleading.

Validation Strategies. Selecting fitting strategy involves deep understanding of the data. It normally involves investigating the variability of model results so that we can distinguish a "luck" from "an actual improvement". However, in medical imaging, datasets are often limited in size, making this method susceptible to high variance in performance estimates.

Cross-Validation. **K-fold cross-validation (CV)** divides the dataset into k subsets (folds). The model is trained on k-1 folds and validated on the remaining fold, a process repeated k times. The results are averaged to produce a performance estimate [79].

Stratified K-fold ensures that each fold maintains the same class distribution as the original dataset, which is crucial for imbalanced datasets common in skin cancer classification. Mahesh et al. (2023) employed stratified K-Fold CV to handle class imbalance problem in [80].

Nested cross-validation addresses the bias in performance estimation due to hyperparameter tuning by having an inner loop for model selection and an outer loop for model assessment.

External validation. This approach involves testing the model on entirely independent datasets from different institutions or populations. This approach provides a robust assessment of the model's generalizability [7]. Brinker et al. (2019) performed an external validation of a DL model for melanoma detection across different populations, emphasizing the necessity of external validation for assessing generalizability [81].

Evaluation Metrics. Accuracy measures the proportion of correct predictions but can be misleading with imbalanced datasets. Precision shows how many identified class samples were identified correctly. Recall (Sensitivity) is how many samples of a class the model was able to identify. Specificity measures the proportion of true negatives among all actual negatives

F1-Score is harmonic mean of precision and recall:

$$F1 = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}}.$$

Balanced accuracy accounts for class imbalance by averaging the recall obtained on each class:

$$Balanced\ Accuracy = \frac{Precision + Recall}{2}.$$

These metrics provide a balanced view of performance, especially important in medical diagnosis where false negatives and false positives have different implications. They are recommended to be used together to get a different perspectives in performance.

Another set of effective metrics is Receiver Operating Characteristic (ROC) Curve and Area Under the Curve (AUC). Receiver Operating Characteristic is a graph with a y-axis representing Sensitivity and an x-axis representing 1 – Specificity. It represents relation of classification threshold and correctly classified samples. In order to easily compare different models AUC is employed — an area under ROC Curve.

Han et al. (2020) utilized ensemble learning and evaluated their model using ROC-AUC, precision-recall curves, and conducted statistical significance testing to demonstrate improvements over previous methods [82].

OPPORTUNITIES FOR INNOVATION

In the domain of DL-based skin cancer classification, there are several promising avenues for innovation that can significantly enhance diagnostic accuracy, patient outcomes, and system scalability. As the field continues to mature, the integration of cutting-edge technologies can address existing challenges, optimize model performance, and broaden access to dermatological diagnostics.

Multimodal Data Integration. Integrating clinical data (e.g., patient demographics, lesion history, and genetic markers) with imaging data has the potential to improve diagnostic accuracy. While current models primarily rely on dermoscopic images, including non-visual patient information can provide additional context, leading to more accurate predictions. For instance, factors like age, lesion location, and personal/family history of skin cancer are crucial in assessing melanoma risk. Models that combine imaging with clinical data have demonstrated improved sensitivity and specificity, particularly for complex cases [83]. Moldovanu et al. (2021) used surface fractal dimensions and statistical color cluster features to improve model quality. Future research should focus on developing architectures capable of effectively fusing diverse data types. Nikitin et al. utilised fractal dimension with ViT focusing on different ways of integration of the metric [84]. Future research should focus on developing architectures capable of effectively fusing diverse data types.

Explainable AI. Building trust in AI-driven diagnostics is essential for clinical adoption. Techniques like Gradient-weighted Class Activation Mapping (Grad-CAM) [85] and Layer-wise Relevance Propagation (LRP) [86] offer explainability by highlighting regions in dermoscopic images that contributed to the model's decision. These visual explanations can help clinicians understand the model's reasoning, enabling them to verify the output and diagnose more confidently. Explainable AI can also aid in identifying potential biases in the model, particularly concerning underrepresented skin tones, thereby addressing disparities in diagnostic outcomes. Future work should explore enhancing the interpretability of DL models while maintaining high classification accuracy, especially in high-stakes settings like oncology.

Edge Computing. Deploying lightweight DL models on edge devices, such as smartphones, can facilitate real-time skin cancer detection, especially in resource-constrained settings. Advances in model optimization, such as pruning, quantization, and using architectures like MobileNet and EfficientNet, can reduce the computational load while maintaining accuracy. This is particularly relevant for underserved regions where access to dermatologists is limited. Edge-based AI systems can provide preliminary assessments, encouraging users to seek medical consultation if a lesion is flagged as suspicious. Research in this area should prioritize developing robust, efficient models that can handle diverse image quality and environmental conditions typical of mobile device usage.

GUIDANCE FOR EMERGING RESEARCHERS

The field of DL for skin cancer classification is both challenging and rewarding, offering ample opportunities for innovation. However, newcomers to the field may face a steep learning curve due to the complexity of data, algorithms, and clinical considerations. Here are some practical tips for emerging researchers to navigate this evolving landscape:

Guidance for Emerging Researchers. The field of DL for skin cancer classification is both challenging and rewarding, offering ample opportunities for innovation. However, newcomers to the field may face a steep learning curve due to the complexity of data, algorithms, and clinical considerations. Here are some practical tips for emerging researchers to navigate this evolving landscape:

1. **Focus on Data Quality and Preprocessing**. One of the biggest hurdles in dermatological AI research is access to high-quality, annotated datasets. Begin by familiarizing yourself with widely used datasets like ISIC and HAM10000. Pay special attention to data preprocessing, including normalization, augmentation, and segmentation techniques, to optimize model performance. Addressing challenges like class imbalance and image noise is crucial for achieving reliable results. However, the real experiments to achieve top metrics results must be conducted on larger data since it is available — as of now ISIC Archive includes more than 400,000 images available to use.

2. **Start with Transfer Learning and Fine-Tunin**g. Given the limited availability of labeled medical data (ISIC Archive is not nearly as big as ImageNet), leveraging transfer learning from pre-trained models on large-scale datasets like ImageNet can accelerate progress. Fine-tuning these models on skin lesion datasets can yield competitive results with relatively less training time. Explore architectures such as ResNet, EfficientNet, and ViT to identify what works best for your specific use case. Also, empirical results show that ensemble models do great in the task.

3. **Embrace Explainability from the Start**. Building interpretability into your models is essential, especially in healthcare applications where clinicians require transparent and understandable outputs. Experiment with tools like Grad-CAM and SHAP [87] to visualize your model's decision-making process. Prioritizing explainability will not only help you gain the trust of the medical community but also ensure that your models can be safely deployed in clinical settings.

4. **Keep Ethics and Privacy at the Forefront**. When dealing with sensitive medical data, ethical considerations and compliance with regulations like GDPR are paramount. Consider approaches like federated learning and differential privacy to ensure patient confidentiality. Being aware of these considerations early on will help you design ethically responsible research projects that can be more easily translated into real-world applications.

5. **Stay Updated on Emerging Trends**. The field of AI in healthcare is rapidly evolving, with new techniques like self-supervised learning, multimodal models, and quantum ML showing potential. Regularly reviewing the latest research, participating in conferences, and engaging with collaborative research communities like the ISIC Challenge can keep you at the forefront of the field. Additionally, exploring adjacent fields like personalized medicine and bioinformatics can open up interdisciplinary opportunities.

By focusing on these areas, emerging researchers can build a strong foundation and contribute meaningfully to the development of AI-driven skin cancer diagnostics.

CONCLUSION

DL has revolutionized the field of skin cancer diagnostics, offering tools that can potentially match, or even surpass, dermatologist-level performance. However, the journey from research to clinical application is fraught with challenges. Our analysis highlights the importance of high-quality data, rigorous preprocessing, and robust model evaluation in developing reliable diagnostic systems. Addressing issues like data privacy, class imbalance, and underrepresentation of diverse skin tones remains critical to ensuring equitable healthcare outcomes.

The integration of clinical metadata with imaging data, along with techniques such as federated learning and edge computing, presents promising avenues to enhance model performance and broaden access to diagnostics, particularly in resource-constrained regions. Additionally, incorporating explainable AI methodologies can help gain the trust of clinicians, paving the way for real-world adoption.

For emerging researchers, focusing on ethical considerations, leveraging transfer learning, and embracing advancements in multimodal integration are vital steps toward impactful contributions in this domain. As AI continues to evolve, its role in skin cancer diagnostics will likely expand, enabling more accurate, accessible, and personalized care. Future research should aim at bridging the gap between technological capabilities and clinical needs, ultimately transforming the landscape of dermatological diagnostics and improving patient outcomes.

REFERENCES

1. *Cancer Research UK*. Available: https://www.cancerresearchuk.org/
2. *Health workforce: The health workforce crisis*. Available: https://www.who.int/news-room/questions-and-answers/item/q-a-on-the-health-workforce-crisis
3. L.G. Wich et al., "Impact of Socioeconomic Status and Sociodemographic Factors on Melanoma Presentation Among Ethnic Minorities," *Journal of Community Health*, vol. 36, no. 3, pp. 461–468, Nov. 2010. doi: 10.1007/s10900-010-9328-4.
4. A. Esteva et al., "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, Jan. 2017. doi: 10.1038/nature21056.
5. *ISIC International Skin Imaging Collaboration*. Available: https://www.isic-archive.com/
6. "Can artificial intelligence detect melanoma?" *Memorial Sloan Kettering Cancer Center*, Jul. 24, 2024. Available: https://www.mskcc.org/news/can-artificial-intelligence-detect-melanoma
7. Y. Liu et al., "A deep learning system for differential diagnosis of skin diseases," *Nature Medicine*, vol. 26, no. 6, pp. 900–908, May 2020. doi: 10.1038/s41591-020-0842-3.
8. A. Yilmaz, M. Kalebasi, Y. Samoylenko, M.E. Guvenilir, H. Uvet, "Benchmarking of Lightweight Deep Learning Architectures for Skin Cancer Classification using ISIC 2017 Dataset," *arXiv* (*Cornell University*), Jan. 2021. doi: 10.48550/arxiv.2110.12270.
9. M. Baygin, T. Tuncer, S. Dogan, "New pyramidal hybrid textural and deep features based automatic skin cancer classification model: Ensemble DarkNet and textural feature extractor," *arXiv (Cornell University)*, Jan. 2022. doi: 10.48550/arxiv.2203.15090.
10. K. Agarwal, T. Singh, "Classification of Skin Cancer Images using Convolutional Neural Networks," *arXiv (Cornell University)*, Jan. 2022. doi: 10.48550/arxiv.2202.00678.
11. D. Wen et al., "Characteristics of publicly available skin cancer image datasets: a systematic review," *The Lancet Digital Health*, vol. 4, no. 1, pp. e64–e74, Nov. 2021. doi: 10.1016/s2589-7500(21)00252-1.
12. "ISIC 2024 — Skin Cancer Detection with 3D-TBP," *Kaggle*. Available: https://www.kaggle.com/competitions/isic-2024-challenge
13. P. Tschandl, C. Rosendahl, H. Kittler, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Scientific Data*, vol. 5, no. 1, Aug. 2018. doi: 10.1038/sdata.2018.161.
14. A.T. Priyeshkumar, G. Shyamala, T. Vasanth, Selvan V. Ponniyin, "Transforming Skin Cancer Diagnosis: A Deep Learning Approach with the Ham10000 Dataset," *Cancer Investigation*, vol. 42, no. 10, pp. 801–814, Nov. 2024. doi: 10.1080/07357907.2024.2422602.
15. T.M. Alam et al., "An efficient Deep Learning-Based skin cancer classifier for an imbalanced dataset," *Diagnostics*, vol. 12, no. 9, p. 2115, Aug. 2022. doi: 10.3390/diagnostics12092115.
16. H.A. Owida et al., "Improved deep learning architecture for skin cancer classification," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 36, no. 1, p. 501, Jul. 2024. doi: 10.11591/ijeecs.v36.i1.pp501-508.
17. A. Ameri, "A deep learning approach to skin cancer detection in dermoscopy images," *Journal of Biomedical Physics and Engineering*, vol. 10, no. 6, Nov. 2020. doi: 10.31661/jbpe.v0i0.2004-1107.

18. S.S. Chaturvedi, J.V. Tembhurne, T. Diwan, "A multi-class skin Cancer classification using deep convolutional neural networks," *Multimedia Tools and Applications*, vol. 79, no. 39–40, pp. 28477–28498, Aug. 2020. doi: 10.1007/s11042-020-09388-2.
19. *ADDI - Automatic computer-based Diagnosis system for Dermoscopy Images*. Available: https://www.fc.up.pt/addi/ph2%20database.html)
20. I. Giotis, N. Molders, S. Land, M. Biehl, M.F. Jonkman, N. Petkov, "MED-NODE: A computer-assisted melanoma diagnosis system using non-dermoscopic images," *Expert Systems With Applications*, vol. 42, no. 19, pp. 6578–6585, May 2015. doi: 10.1016/j.eswa.2015.04.034.
21. Dermofit Image Library available from The University of Edinburgh. Available: https://licensing.edinburgh-innovations.ed.ac.uk/product/dermofit-image-library
22. X. Sun, J. Yang, M. Sun, K. Wang, "A benchmark for automatic visual classification of clinical skin disease images," in *Lecture notes in computer science*, 2016, pp. 206–222. doi: 10.1007/978-3-319-46466-4_13.
23. "Datasets used to train AI to detect skin cancer lack information on darker skin and often incomplete," *Haiku*, Nov. 10, 2021. Available: https://www.cancer.ox.ac.uk/ news/lack-of-data-available-to-detect-skin-cancer-in-darker-skin
24. S. Hu, R.M. Soza-Vento, D.F. Parker, R.S. Kirsner, "Comparison of stage at diagnosis of melanoma among Hispanic, Black, and white patients in Miami-Dade County, Florida," *Archives of Dermatology*, vol. 142, no. 6, Jun. 2006. doi: 10.1001/archderm.142.6.704.
25. H.M. Gloster, K. Neal, "Skin cancer in skin of color," *Journal of the American Academy of Dermatology*, vol. 55, no. 5, pp. 741–760, Oct. 2006. doi: 10.1016/j.jaad.2005.08.063.
26. "Health Insurance Portability and Accountability Act (HIPAA)", *U.S. Department of Health and Human Services*. Available: https://www.hhs.gov/hipaa/index.html
27. "General Data Protection Regulation (GDPR) – Legal Text," *General Data Protection Regulation (GDPR)*, Apr. 22, 2024. Available: https://gdpr-info.eu/
28. S. Khan, H. Ali, Z. Shah, "Identifying the role of vision transformer for skin cancer — A scoping review," *Frontiers in Artificial Intelligence*, vol. 6, Jul. 2023. doi: 10.3389/frai.2023.1202990.
29. P. Tschandl et al., "Human–computer collaboration for skin cancer recognition," *Nature Medicine*, vol. 26, no. 8, pp. 1229–1234, Jun. 2020. doi: 10.1038/s41591-020-0942-0.
30. T.J. Brinker et al., "Deep neural networks are superior to dermatologists in melanoma image classification," *European Journal of Cancer*, vol. 119, pp. 11–17, Aug. 2019. doi: 10.1016/j.ejca.2019.05.023.
31. S.F.H. Shah, D. Arecco, H. Draper, S. Tiribelli, E. Harriss, R.N. Matin, "Ethical implications of artificial intelligence in skin cancer diagnostics: use-case analyses," *British Journal of Dermatology*, Nov. 2024. doi: 10.1093/bjd/ljae434.
32. E.R. Gordon et al., "Ethical considerations for artificial intelligence in dermatology: a scoping review," *British Journal of Dermatology*, vol. 190, no. 6, pp. 789–797, Feb. 2024. doi: 10.1093/bjd/ljae040.
33. H.A. Haenssle et al., "Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists," *Annals of Oncology*, vol. 29, no. 8, pp. 1836–1842, May 2018. doi: 10.1093/annonc/mdy166.
34. A.G.C. Pacheco, R.A. Krohling, "Recent advances in deep learning applied to skin cancer detection," *arXiv (Cornell University)*, Jan. 2019. doi: 10.48550/arxiv.1912.03280.
35. P. Popecki, M. Kozakiewicz, M. Ziętek, K. Jurczyszyn, "Fractal Dimension Analysis of Melanocytic nevi and melanomas in Normal and Polarized Light—A Preliminary Report," *Life*, vol. 12, no. 7, p. 1008, Jul. 2022. doi: 10.3390/life12071008.
36. "Melanoma Skin Cancer Dataset of 10000 Images," *Kaggle*, Mar. 29, 2022. Available: https://www.kaggle.com/datasets/hasnainjaved/melanoma-skin-cancer-dataset-of-10000-images
37. S. Shen et al., "A Low-Cost High-Performance data augmentation for Deep Learning-Based skin lesion classification," *BME Frontiers*, 2022. doi: https://doi.org/ 10.34133/2022/9765307
38. G.M.S. Himel, Md.M. Islam, Kh.A. Al-Aff, S.I. Karim, Md.K.U. Sikder, "Skin Cancer Segmentation and Classification Using Vision Transformer for Automatic Analysis in Dermatoscopy-Based Noninvasive Digital System," *International Journal of Biomedical Imaging*, vol. 2024, pp. 1–18, Feb. 2024. doi: 10.1155/2024/3022192.

39. K. Polat, K.O. Koc, "Detection of Skin Diseases from Dermoscopy Image Using the combination of Convolutional Neural Network and One-versus-All," *Journal of Artificial Intelligence and Systems*, vol. 2, no. 1, pp. 80–97, Jan. 2020. doi: 10.33969/ais. 2020.21006.
40. B. Walker et al., "Dermoscopy diagnosis of cancerous lesions utilizing dual deep learning algorithms via visual and audio (sonification) outputs: Laboratory and prospective observational studies," *EBioMedicine*, vol. 40, pp. 176–183, 2019. doi: https://doi.org/10.1016/j.ebiom.2019.01.028
41. M.S. Ali, M.S. Miah, J. Haque, M.M. Rahman, M.K. Islam, "An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models," *With Applications*, vol. 5, 100036, 2021. doi: https://doi.org/ 10.1016/j.mlwa.2021.100036
42. Y. Wu, B. Chen, A. Zeng, D. Pan, R. Wang, S. Zhao, "Skin Cancer Classification With Deep Learning: A Systematic Review," *Frontiers in Oncology*, vol. 12, Jul. 2022. doi: 10.3389/fonc.2022.893972.
43. D. Bisla, A. Choromanska, R.S. Berman, J.A. Stein, D. Polsky, "Towards Automated Melanoma Detection With Deep Learning: Data Purification and Augmentation," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Jun. 2019*. doi: 10.1109/cvprw.2019.00330.
44. Z. Qin, Z. Liu, P. Zhu, Y. Xue, "A GAN-based image synthesis method for skin lesion classification," *Computer Methods and Programs in Biomedicine*, vol. 195, p. 105568, May 2020. doi: 10.1016/j.cmpb.2020.105568.
45. B. Ahmad, S. Jun, V. Palade, Q. You, L. Mao, M. Zhongjie, "Improving Skin Cancer Classification Using Heavy-Tailed Student T-Distribution in Generative Adversarial Networks (TED-GAN)," *Diagnostics*, vol. 11, no. 11, p. 2147, Nov. 2021. doi: 10.3390/diagnostics11112147.
46. I.S.A. Abdelhalim, M.F. Mohamed, Y.B. Mahdy, "Data augmentation for skin lesion using self-attention based progressive generative adversarial network," *Expert Systems With Applications*, vol. 165, p. 113922, Sep. 2020. doi: 10.1016/j.eswa.2020.113922.
47. R. Kaur, H. Gholam Hosseini, R. Sinha, "Synthetic Images Generation Using Conditional Generative Adversarial Network for Skin Cancer Classification," *TENCON 2021 - 2021 IEEE Region 10 Conference (TENCON)*, vol. 27, pp. 381–386, Dec. 2021. doi: 10.1109/tencon54134.2021.9707291.
48. J. Pope, M. Hassanuzzaman, M. Sherpa, O. Emara, A. Joshi, N. Adhikari, "Skin Cancer Machine Learning Model Tone Bias," *arXiv (Cornell University)*, 2024. doi: https://doi.org/10.48550/arxiv.2410.06385
49. E. Rezk, M. Eltorki, W. El-Dakhakhni, "Improving Skin Color Diversity in Cancer Detection: Deep Learning Approach," *JMIR Dermatology*, 5(3), e39143, 2022. doi: https://doi.org/10.2196/39143
50. B.N. Walker et al., "Skin cancer detection in diverse skin tones by machine learning combining audio and visual convolutional neural networks," *Oncology*, 103(5), pp. 113–120, 2024. doi: https://doi.org/10.1159/000541573
51. J. Han, M. Kamber, J. Pei, *Data Mining: Concepts and Techniques*; 3rd edition, 2011. Available: https://www.scholartext.com/book/88809627?_locale=fr
52. T. Majtner, B. Bajić, S. Yildirim, J.Y. Hardeberg, J. Lindblad, N. Sladoje, "Ensemble of Convolutional Neural Networks for Dermoscopic Images Classification," *arXiv (Cornell University)*, 2018. doi: https://doi.org/10.48550/arxiv.1808.05071
53. S.M. Pizer et al., "Adaptive histogram equalization and its variations," *Computer Vision, Graphics, and Image Processing*, 39(3), pp. 355–368, 1987.
54. K. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics gems IV*. Academic Press Professional, Inc., 1994, pp. 474–485.
55. C.M. Bishop, *Pattern recognition and machine learning*. Springer, 2006.
56. V.D. Midasala, B. Prabhakar, J.K. Chaitanya, K. Sirnivas, D. Eshwar, P.M. Kumar, "MFEUsLNet: Skin cancer detection and classification using integrated AI with multilevel feature extraction-based unsupervised learning," *Engineering Science and Technology an International Journal*, 51(10), 101632, 2024. doi: https://doi.org/ 10.1016/j.jestch.2024.101632
57. T.K. Lee et al., "Dullrazor®: A software approach to hair removal from images," *Computers in Biology and Medicine*, 27(6), pp. 533–543, 1997.
58. R. Maurya et al., "Skin cancer detection through attention guided dual autoencoder approach with extreme learning machine," *Scientific Reports*, vol. 14, no. 17785, Aug. 2024. doi: 10.1038/s41598-024-68749-1.

59. A. Kirillov et al., “Segment Anything,” *arXiv (Cornell University)*, Apr. 2023. doi: https://doi.org/10.48550/arxiv.2304.02643
60. A. Dosovitskiy et al., “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale,” *arXiv.org*, Jun. 03, 2021. Available: https://www.arxiv.org/abs/2010.11929v2
61. S. Targ, D. Almeida, K. Lyman, “Resnet in Resnet: Generalizing Residual Architectures,” *arXiv (Cornell University)*, Jan. 2016ю doi: 10.48550/arxiv.1603.08029.
62. X. Han, Y. Zhong, L. Cao, L. Zhang, “Pre-Trained AlexNet Architecture with Pyramid Pooling and Supervision for High Spatial Resolution Remote Sensing Image Scene Classification,” *Remote Sensing*, vol. 9, no. 8, p. 848, Aug. 2017. doi: 10.3390/rs9080848.
63. Q. Guan et al., “Deep convolutional neural network VGG-16 model for differential diagnosing of papillary thyroid carcinomas in cytological images: a pilot study,” *Journal of Cancer*, vol. 10, no. 20, pp. 4876–4882, Jan. 2019. doi: 10.7150/jca.28769.
64. P. Carcagnì et al., “Classification of Skin Lesions by Combining Multilevel Learnings in a DenseNet Architecture,” in *Lecture notes in computer science*, pp. 335–344, 2019. doi: 10.1007/978-3-030-30642-7_30.
65. D. Sinha, M. El-Sharkawy, “Thin MobileNet: An Enhanced MobileNet Architecture,” *IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, pp. 0280–0285, Oct. 2019. doi: 10.1109/uemcon47517. 2019.8993089.
66. L.N. Smith, “Cyclical learning rates for training neural networks,” in *2017 IEEE Winter Conference on Applications of Computer Visio*n, pp. 464–472.
67. I. Loshchilov, F. Hutter, “SGDR: Stochastic gradient descent with warm restarts,” *arXiv preprint*, 2016. doi: https://doi.org/10.48550/arXiv.1608.03983
68. D.P. Kingma, J. Ba, “Adam: A method for stochastic optimization,” in *3rd International Conference on Learning Representations (ICLR)*, 2015.
69. R. Ali, A. Manikandan, R. Lei, J. Xu, “A novel SpaSA based hyper-parameter optimized FCEDN with adaptive CNN classification for skin cancer detection,” *Scientific Reports*, vol. 14, no. 1, Apr. 2024. doi: 10.1038/s41598-024-57393-4.
70. Duyen N.T. Le, Hieu X. Le, Lua T. Ngo, Hoan T. Ngo, “Transfer learning with class-weighted and focal loss function for automatic skin cancer classification,” *arXiv (Cornell University)*, Jan. 2020. doi: https://doi.org/10.48550/arxiv.2009.05977
71. S. Ghosh, S. Dhar, Raktim Yoddha, S. Kumar, Abhinav Kumar Thakur, Nanda Dulal Jana, “Melanoma Skin Cancer Detection Using Ensemble of Machine Learning Models Considering Deep Feature Embeddings,” *Procedia Computer Science*, vol. 235, pp. 3007–3015, Jan. 2024. doi: https://doi.org/10.1016/j.procs.2024.04.284
72. J. Atwood, D. Towsley, “Diffusion-Convolutional Neural Networks,” *arXiv.org*, Jul. 08, 2016. Available: https://arxiv.org/abs/1511.02136
73. S. Sabour, N. Frosst, G.E. Hinton, “Dynamic Routing Between Capsules,” *arXiv (Cornell University)*, Jan. 2017. doi: https://doi.org/10.48550/arxiv.1710.09829
74. “SIIM-ISIC Melanoma Classification,” *Kaggle*, 2024. Accessed on: Dec. 28, 2024. Available: https://www.kaggle.com/competitions/siim-isic-melanoma-classification/ discussion/181830
75. K. Chaitanya, E. Erdil, N. Karani, E. Konukoglu “Contrastive learning of global and local features for medical image segmentation with limited annotations,” in *Advances in Neural Information Processing Systems*, 2020.
76. T. Chen, S. Kornblith, M. Norouzi, G. Hinton, “A simple framework for contrastive learning of visual representations,” in *International Conference on Machine Learning*, 2020.
77. S. Azizi et al., “Big self-supervised models advance medical image classification,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021.
78. H. Haggerty, R. Chandra, “Self-supervised learning for skin cancer diagnosis with limited training data,” *arXiv (Cornell University)*, Jan. 2024. doi: https://doi.org/10.48550/arxiv.2401.00692
79. P. Refaeilzadeh, L. Tang, H. Liu, “Cross-validation,” in *Encyclopedia of Database Systems*, pp. 532–538. Springer, 2009.
80. T.R. Mahesh, V.V. Kumar, D.V. Kumar, O. Geman, M. Margala, M. Guduri, “The stratified K-folds cross-validation and class-balancing methods with high-performance ensemble classifiers for breast cancer classification,” *Healthcare Analytics*, vol. 4, p. 100247, Sep. 2023. doi: 10.1016/j.health.2023.100247.

81. T.J. Brinker et al., "Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task," *European Journal of Cancer*, vol. 113, pp. 47–54, 2019.
82. S.S. Han et al., "Deep neural networks show an equivalent and often superior performance to dermatologists in onychomycosis diagnosis: Automatic construction of onychomycosis datasets by region-based convolutional deep neural network," *PLOS One*, vol. 13, no. 1, p. e0191493, Jan. 2018. doi: https://doi.org/10.1371/ journal.pone.0191493
83. S. Moldovanu, F.A. Damian Michis, K.C. Biswas, A. Culea-Florescu, L. Moraru, "Skin Lesion Classification Based on Surface Fractal Dimensions and Statistical Color Cluster Features Using an Ensemble of Machine Learning Techniques," *Cancers*, vol. 13, no. 21, p. 5256, Oct. 2021. doi: https://doi.org/10.3390/cancers13215256
84. V. Nikitin, V. Danylov, "Integration of Fractal Dimension in Vision Transformer for Skin Cancer Classification," *Electronics and Control Systems*, vol. 2, no. 76, pp. 15–20, Jun. 2023. doi: https://doi.org/10.18372/1990-5548.76.17662
85. R.R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization," *International Journal of Computer Vision*, vol. 128, no. 2, pp. 336–359, Feb. 2020. doi: https://doi.org/10.1007/s11263-019-01228-7
86. A. Binder, G. Montavon, S. Bach, K.-R. Müller, and W. Samek, "Layer-wise Relevance Propagation for Neural Networks with Local Renormalization Layers," *arXiv (Cornell University)*, Apr. 2016. doi: https://doi.org/10.48550/arxiv.1604.00825
87. S. Lundberg, S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," *arXiv.org*, Nov. 24, 2017. Available: https://arxiv.org/abs/1705.07874v2

Received 09.01.2025

INFORMATION ON THE ARTICLE

Vladyslav O. Nikitin, ORCID: 0009-0001-9921-0213, Educational and Research Institute for Applied System Analysis of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine, e mail: nvo63911@gmail.com

Valery Ya. Danilov, ORCID: 0000-0003-3389-3661, Educational and Research Institute for Applied System Analysis of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine, e mail: danilov1950@ukr.net

ПОДОЛАННЯ ВИКЛИКІВ У ГЛИБОКОМУ НАВЧАННІ ДЛЯ ВИЯВЛЕННЯ РАКУ ШКІРИ / В.О. Нікітін, В.Я. Данилов

Анотація. Рак шкіри є одним із найпоширеніших злоякісних новоутворень у світі. Раннє виявлення є критичним фактором зниження рівня смертності. Це підкреслює необхідність доступних систем комп'ютерної діагностики. Нещодавні досягнення в глибокому навчанні показали великі перспективи у вирішенні цієї проблеми. Незважаючи на цей прогрес у галузі машинного навчання, дослідники стикаються із численними перешкодами, коли йдеться про класифікацію раку шкіри. Розглянуто сучасний стан діагностики раку шкіри на основі глибокого навчання, критичні аспекти розроблення системи, включно з попереднім обробленням даних, навчанням моделей та оцінкою продуктивності. Крім того, висвітлюються можливості для інновацій, які можуть значно просунути цю галузь. Надаючи вичерпний огляд, стаття має на меті допомогти дослідникам та практикам в оптимізації моделей глибокого навчання, усуненні існуючих обмежень та дослідженні нових тенденцій для підвищення точності та доступності діагностики.

Ключові слова: рак шкіри, глибоке навчання, класифікація, трансформатори, CNN, GAN, попереднє оброблення даних, доповнення даних.

UDC 004.6:33
DOI: 10.20535/SRIT.2308-8893.2025.2.04

A CONCEPTUAL MODEL AND A SYSTEM FOR REPLACING TEXT IN AN IMAGE WHILE PRESERVING THE STYLE

P. MASLIANKO, M. ROMANOV

Abstract. Text replacement in images, particularly while preserving its style, is a complex task that requires solving a range of scientific challenges and developing new technical solutions. One of the main issues is maintaining the authenticity and harmony of the image after modifications. The Research Objective is the development of a conceptual model and a system for text replacement in images with style preservation based on systems engineering methodology and the Eriksson-Penker business profile, ensuring the natural integration of new text elements into the image's context. Implementation Methodology – the systems engineering methodology and the Eriksson-Penker business profile are used to formalize the structured process of developing a system for text replacement in images with style preservation. Research Results – a method for developing the system based on systems engineering techniques was proposed, consisting of four main stages. In the first stage, the system structure is modeled as an Eriksson-Penker business profile. In the second stage, a set of processes is defined that are characteristic of the Data Science system class and the CRISP-DM international standard. Also, the structural and dynamic representations of the conceptual model, as well as the component interaction interfaces, are modeled. The third stage involves implementing a specific version of the system, while the fourth stage focuses on system verification and validation. A systems engineering method for the conceptual model and system for text replacement in images with style preservation has been proposed. It is based on a modified Eriksson-Penker business profile for metalevel system representation and international standards for Data Science and Data Mining processes.

Keywords: systems engineering method, Eriksson-Penker business profile, conceptual model, system for text replacement in images with style preservation.

BACKGROUND

Replacing text in an image, in particular, while preserving its style, is a complex task that requires solving a number of scientific problems and new technical solutions. One of the main challenges is to preserve the authenticity and harmony of the image after the changes have been made, as any modification of the text can disrupt the original style, fonts, colours and visual composition. To make such changes look natural, a mechanism is needed that can effectively identify the text style, recreate it, and adapt it to the new content.

Key aspects of the task:

1. Preserving the style: when replacing text, it is important not only to correctly reproduce the font, size, colour and other visual characteristics, but also to integrate it into the overall visual aesthetics of the image (including background, shadows, textures).

2. Adaptability to different styles: texts on images can be executed in a wide variety of styles — from handwritten text to complex graphic elements. The sys-

tem should be able to work with different text styles and formats, automatically detecting them.

3. Contextual replacement: Often, text is not just a separate element, but is integrated into a complex visual scene. Therefore, the replacement should take into account not only the style but also the context of the image to make the new text look appropriate.

4. Automation and scalability: manual text replacement is a time-consuming process in many cases. Automation of this task will significantly save time with a large volume of images, which is important for industries such as marketing, design or content localisation (Table 1).

Table 1. Overview and comparative analysis of existing solutions

Approach	Advantages	Disadvantages
Manual tools	High precision at the right skill level, full user control	Labour intensive, difficult to scale
Classic computer vision algorithms	Speed, automation of key processes	Limited ability to save complex text styles, need for additional customisation
Generative models (GAN)	Ability to automatically save complex stylistic elements of text	High cost of computing, dependence on data quality, long learning curve

Current methods for style-preserving text replacement in images are mostly based on generative adversarial networks and deep learning. StyleGAN remains one of the most effective solutions, thanks to its ability to accurately reproduce stylistic elements of text and integrate them into the overall context of the image.

Motivation for the development of the conceptual model and system to replace text in an image while preserving the style.

The main motivation is the need for a system that will automatically replace text in an image while maintaining its style. This is necessary for such cases of activity:

1. Content localisation: Many companies and brands need to adapt their visuals for different markets and languages. For example, advertising banners or posters may have text that needs to be translated and replaced in other languages without losing style and graphic authenticity.

2. Graphic design: Quickly changing text on design mockups can be important when working on prototypes or changes during the approval phase of projects.

Thus, the relevance of the problem lies in the need to develop a scientifically grounded conceptual model of this class of systems and implement an automated tool for replacing text in images while preserving the style. The solution to this problem opens up opportunities to speed up the content creation process, improve its quality and accuracy of visual adaptation for various purposes.

PROBLEM STATEMENT

Object of research. Conceptual ideas and approaches, existing mathematical models and algorithms for detecting, analysing and synthesising text in images, their theoretical aspects, architecture and existing software tools for image and text processing.

Subject of research. Conceptual model and system of text replacement in an image with style preservation. Structural and dynamic representation of the system, image generation algorithms. Models and algorithms for style control in the process of image generation.

Research objective. To develop a conceptual model and a system for replacing text in images while preserving its style, which allows to accurately identify, analyse and reproduce the visual characteristics of the text (font, size, colour, orientation, textures, shadows and other stylistic elements), ensuring the natural integration of new text elements into the context of the image.

Final result. A conceptual model and a system for replacing text in an image while preserving its style, as well as verification and validation results.

A method of system engineering of a conceptual model and a system for replacing text on images while preserving its style

The main idea behind the development of the conceptual model is based on [1–4] and consists in applying the systems engineering methodology and the Eriksson–Penker business profile to formalise an orderly way of developing a text-to-picture image replacement system with style preservation (Fig. 1).

One of the most common models of activity representation is the Eriksson–Penker business profile [5–8], in the context of which the authors formulated four main essences of the formal representation of the activity of any business system (Fig. 1):

– goals (represent the purpose of the system and are formulated as a rule. Goals can be broken down into sub-goals and achieved through the implementation of processes);

– processes (the main actions that make up the system's activities and are intended to achieve the goal in accordance with the established business rules. Processes are usually subject to rules, can change the state of input resources, and produce new resources — the system's output resources — in accordance with the conditions and requirements set by stakeholders);

– resources (physical, abstract or informational objects that the system consumes, uses, processes and produces throughout its activities to achieve the goal);

– rules (certain formalised restrictions, frameworks, conditions and requirements, etc. that are imposed on processes and describe the nature of the relationships between resources).

Such an ordered set of formalised (in particular, in UML notation) entities and system representations based on the Eriksson–Penker business profile formalises and systemises the conceptual model (meta-model) of a text-to-picture image replacement system with style preservation (Fig. 1).

Thus, the essence of the system engineering method of the conceptual model of the image text replacement system with style preservation is to apply the system approach and the Eriksson–Penker business profile to formalise and produce such systems

At the second stage, we formalise the structural and dynamic representation of the conceptual model of the system, interfaces for component interaction, technical requirements and specifications of all stakeholders.

At the third stage, we implement a specific version of the system based on technological and mathematical tools and in accordance with the technical requirements and specifications of all stakeholders.

The fourth stage involves verification and validation of a specific version of the system.

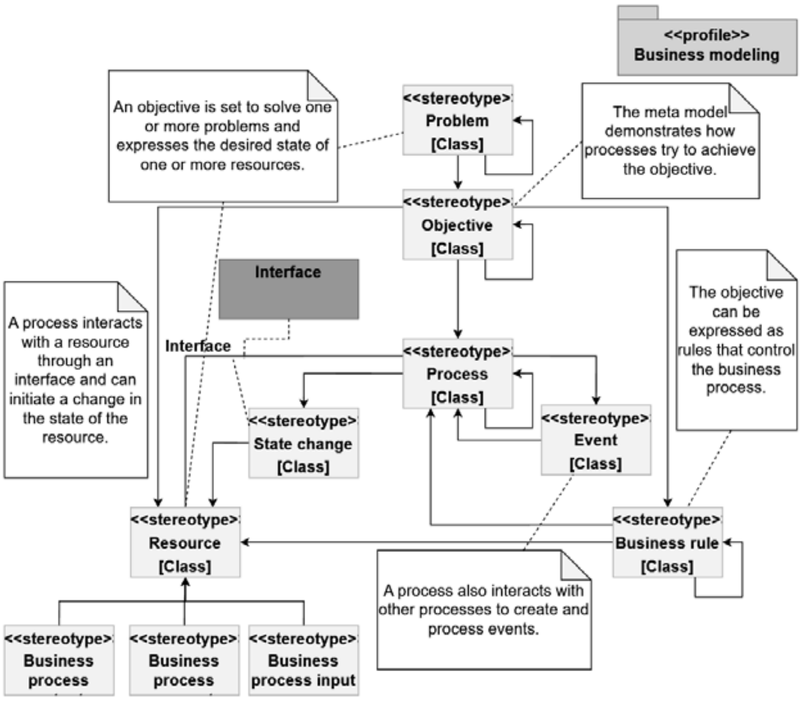


Fig. 1. Improved Eriksson–Penker business profile. Class diagram in UML notation [1]

Formalisation of Eriksson–Penker business profile classes for a conceptual model and a system for replacing text on images while preserving its style

Let's define the content of each of the classes of the diagram (Fig. 1) in terms of the system engineering problem, namely classes:

1. Problem (an actual issue that requires appropriate solutions, the main motivation for developing a system that leads to the formulation of a specific goal. The problem of this work: the need to develop a conceptual model and an automated tool for replacing text in images while preserving the style. The solution to this problem opens up opportunities for significantly speeding up the content creation process, improving its quality and accuracy of visual adaptation for various purposes).

2. Purpose (expresses the global goal of the work designed to solve the problem. The purpose of this work: To develop a system for automatically replacing text in images while preserving its style, which allows for accurate identification, analysis and reproduction of visual text characteristics (font, size, colour, orientation, textures, shadows and other stylistic elements), ensuring natural integration of new text elements into the image context).

3. Process (a set of processes of the system's activity, which results in achieving the goal, a clearly defined sequence of actions/subprocesses that leads to the fulfilment of a certain task. The processes of this system are: Loading training data, Pre-processing training data, Training the model, Replacing text in an image with preservation of style, Functioning of the web application).

4. State change (possible changes in certain resources as a result of the processes. The system has three state changes: Image with text and word coordinate file → Segmented images (the Pre-process training data process), Segmented images with text labels → Transformed images and text labels (the Replace text in image with preservation of style process), Initialised model → Trained model (the Train model process)).

5. Resource (any entities (tangible or intangible) that are consumed and produced by the system under development.

The resources of the lowest level of the hierarchy, which are directly involved in the processes, are also divided into the following three classes by the nature of their influence on the processes:

– Business process output (resources produced by the system, the end result of its operation. This includes the generated image with new text).

– Business process support resource (resources that support the execution of processes, but are not the final result of work: Trained model, Computer hardware and computing resources, Pre-processing algorithms, Metadata files, Software).

– Business process input (primary input resources of the initial processes that initialise the system cycle: Image with text, Metadata file for segmentation, Image for text replacement, New text).

6. Event (occurs due to certain external factors or as a result of interaction between processes. The potential events of this system are Uploading an image with text by a user, Entering a new text to replace it, Loading training data, Completing preprocessing, Completing model training, Transforming input parameters, Outputting results via the interface).

7. Business Rule (BR): formal instructions that regulate, limit, establish the context and framework for the functioning of processes. Example of a business rule: the format of images must be JPG, JPEG, PNG).

Thus, on the basis of the conceptual model (meta-model) of the image text replacement system with style preservation shown in Fig. 1, we can reasonably formalise the structural representation of the image text replacement system with style preservation in the form of a component diagram (Fig. 2).

Formalisation of the structural representation of the conceptual model of the system

Such a conceptual model of structural representation (Fig. 2) formalises the class of systems for replacing text in an image with preserving the style in the form of an ordered set of entities and relations between them. An important property of such a conceptual model is the ability to implement the internal structure of the system components on the basis of various engineering and mathematical tools necessary for the implementation of a particular system without changing the structure and interfaces of interaction between the components.

Here is a short list of the system's interfaces and their functionality:

- IDD (Interface Dataset Download) — interface for downloading an external data set;
- IIP (Interface Image Preprocessing) — interface for transferring the initially processed downloaded data set;
- ISIP (Interface Segmented Image Processing) — interface for transferring processed segmented images with annotations;
- IMT (Interface Model Training) — interface for transferring the trained model;
- IIITP (Interface Input Image and Text Processing) — an interface for transmitting the result of processing images and text entered by the user.

Based on the model of structural representation, we formalise the model of dynamic representation of the system, which shows the internal structure of components and the algorithm of interaction between them (Figs. 3, 4).

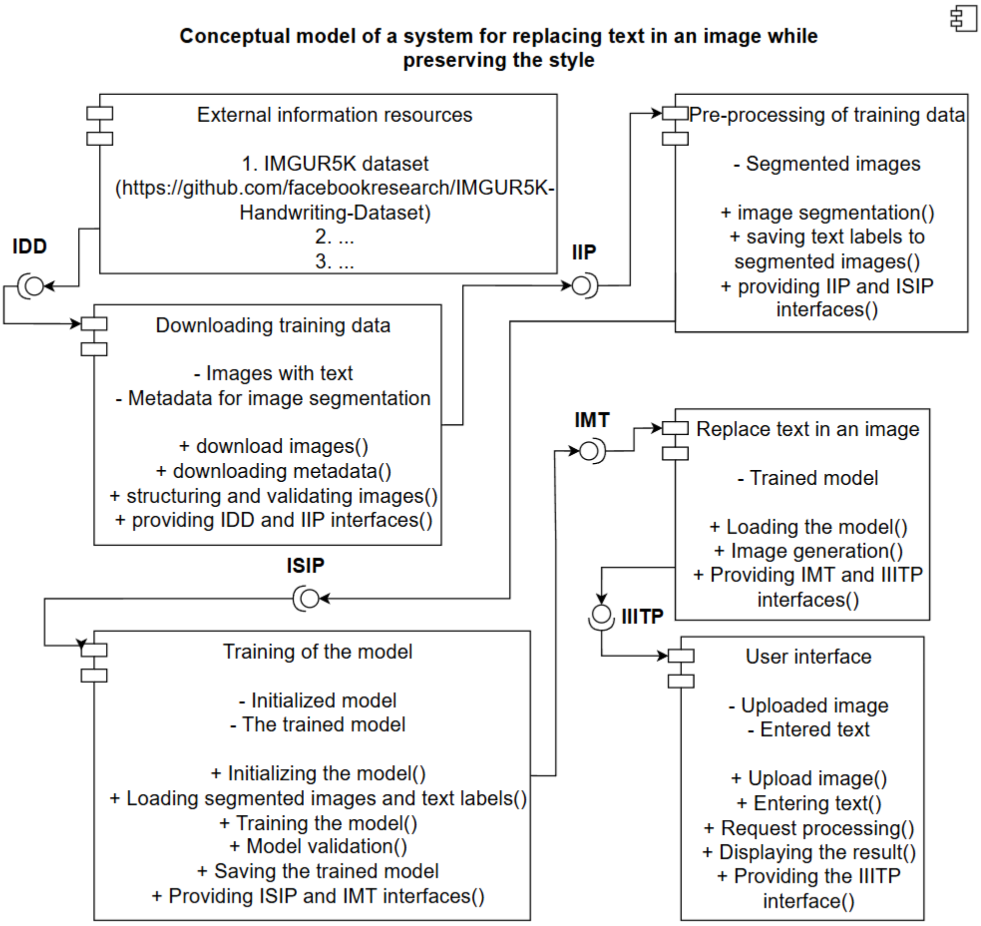


Fig. 2. Conceptual model of a system for replacing text in an image while preserving the style. Component diagram in UML notation

Formalisation of the dynamic representation of the conceptual model of the system

The formalisation of the dynamic representation is determined on the basis of a set of processes specific to the class of Data Science systems, according to the Data Science process defined by O'Neill and Schutt [2] and the international standard CRISP-DM as interpreted by Foster and Fawcett [3] (Figs. 3, 4).

At this level, the organisation of activities can be refined to take into account the specifics of the system and, as a result, decomposed into the following three sub-stages:

1. Collection, analysis, and processing of training text data to be used in model training according to the Data Science process model proposed by O'Neill and Schutt [2] or the Data under standing and Data processing stages of the CRISP-DM (CRoss Industry Standard Process for Data Mining) analysed by Foster and Fawcett [3].

2. The actual construction (architecture development) and training of the model is analogous to the Machine Learning Algorithms Statistical Models stage [2] or the Modeling stage of the standard Data Mining process [3].

3. Determining metrics for evaluating the effectiveness of both trained models and the system as a whole is analogous to the Report Findings stage [2] or the Evaluation stage of the standard Data Mining process [3].

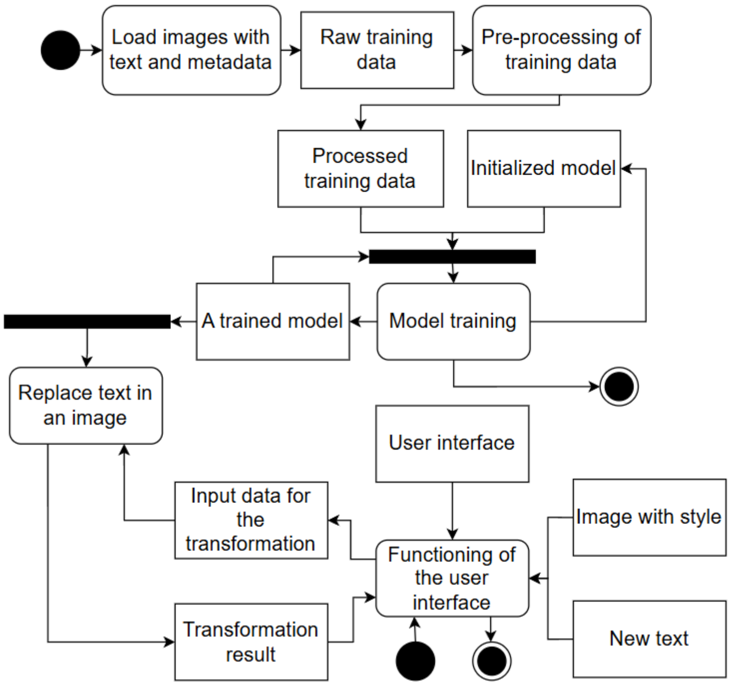


Fig. 3. Conceptual model of a system for replacing text in an image while preserving the style. The first level activity diagram in UML notation

Thus, the Eriksson–Penker business profile is a system of classes and relationships between them that are necessary and sufficient to represent and develop a conceptual model of the system. And the corresponding set of Data Science technologies are tools for implementing the components of the conceptual model of the system.

An exhaustive list of business rules (specifications) for a business profile regulates the functionality of the user interface of a particular system, for example:

BR1 — The user must enter new text and select a replacement image before starting the process. If this data is not provided, the system does not start processing and sends a request to fill in all fields.

BR2 — If the user uploads an image for the style, it must comply with the established formats (JPEG, JPG, PNG) and be no larger than 10 MB.

BR3 — After the text replacement is complete, the user can preview the result before uploading it.

BR4 — If an error occurs (incorrect image loading, segmentation errors, insufficient data for training, etc.), the system should stop the process and send a clear error message to the user indicating possible ways to fix it.

Implementation of the conceptual model for a specific version of the system for replacing text in an image while preserving the style. Implementation of the structural representation of the system

Since the conceptual model of the system is a system of classes and relations between them, necessary and sufficient for the functioning of the system (Fig. 1), and the conceptual model of the system in the form of a diagram of components has the form shown in Fig. 2, then the purpose and functionality of the system components will be as follows:

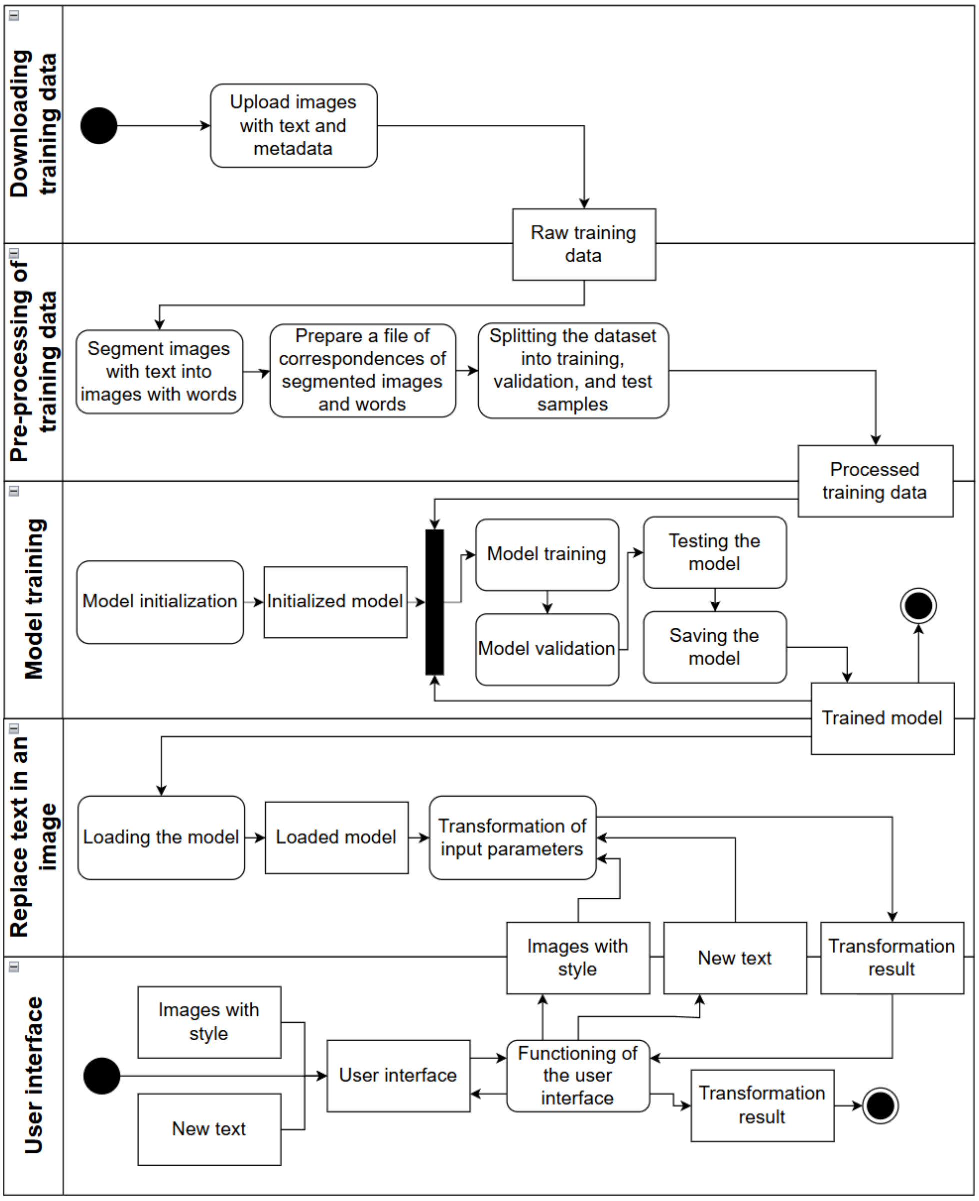


Fig. 4. Conceptual model of a system for replacing text in an image while preserving the style. A detailed process diagram based on a second-level activity diagram in UML notation

1. Training data loading. This module is responsible for loading images from the IMGUR5K dataset [9], checking their integrity using hashes, and loading annotations for use in further processing steps.

2. Training data preprocessing. This module is responsible for segmenting the uploaded images and saving them to the appropriate directories for training, validation and testing. It is also possible to reduce the size of the training dataset by random sampling, which is useful in situations of lack of computing power.

3. Model training. The main module of the system, which results in a trained model for transforming the entered images and text. For training, pre-segmented images are used with corresponding annotations that describe the text on the segmented image.

4. Replacing text with images. A module that serves as a facade for interacting with the trained model, where the trained model is loaded into memory, input parameters are passed to the model, and the results are returned in the form of generated images.

5. User interface. A module that is an interface (web page) where a user can upload their image with a style and enter text to generate an image with text in the style of the uploaded image.

Implementation of a dynamic system representation. The dynamic representation of the system in the form of an activity diagram (Fig. 3) is a visualisation of the main cycles of the processes and their first-level interaction.

Fig. 4 shows a diagram of the processes of the conceptual model of the system in the form of a diagram of interaction of components (see Fig. 2) and shows their internal structure. Such a representation makes it possible to group the main processes of the system by the relevant components. The process diagram is an integral part of the conceptual model based on the Eriksson–Penker business profile, as it combines both types of representations of the system, structural and dynamic, and formalises its activities.

Implementation of the Training Data Loading and Training Data Preprocessing components

Data download module. This module is responsible for loading images from the IMGUR5K dataset [9], checking their integrity using hashes, and loading annotations for use in further processing steps.

Uploading images: For each image, the module generates a URL, uses the requests library to download it, and saves the file in a specific directory. After the image is uploaded, its hash is checked to ensure authenticity.

Parallel uploading: To optimise the uploading process, the module uses the ThreadPoolExecutor from the futures library, which allows you to upload several images simultaneously, which significantly speeds up processing.

Loading annotations: After the images are successfully uploaded, the module loads annotations for each image in JSON format, containing information about the image, its hash, and the coordinates of text elements. The annotations are also divided into three sets: training, validation, and test.

Training data processing module. This module is responsible for processing the uploaded images, extracting text features, and preparing the dataset for model training.

Processing of text elements: For each image, the module uses annotations to find text features in the images. The crop_minAreaRect() function is used to crop the image around the text, taking into account its coordinates and orientation. This function is based on the perspective transformation algorithm cv2.getPerspectiveTransform() from the OpenCV library, which allows you to accurately select text taking into account its slope.

Saving processed images: The cropped text elements are saved in PNG format, and a corresponding record is created in the JSON file with the text on the image for each of them.

Reducing the data sample: If you specify the reduce option, the module randomly selects a part of the images for processing, which allows you to work with a smaller sample for testing or speeding up the work.

Implementation of the Model Training component. Model training process:

1. Load the training dataset (segmented images with annotations).
2. Creating an initialised model of the initial architecture using available python libraries and frameworks / loading a pre-trained model — the latest version of the saved model.
3. Training the model using the GAN algorithm on the training data set by optimising the loss function.
4. Conducting validation to determine the optimal model architecture and model hyperparameters on the validation dataset with the calculation of MSE, SSIM, PSNR and FID metrics.
5. Final testing on the test data set.
6. Saving the current version of the model, selecting the most efficient ve
7. rsion for processing user requests.

Mathematical support of the model training component

In Fig. 5 we can see a high-level diagram of the model training process. This diagram is based on the TextStyleBrush architecture [9]. It shows the main components of the model, and we will describe them:

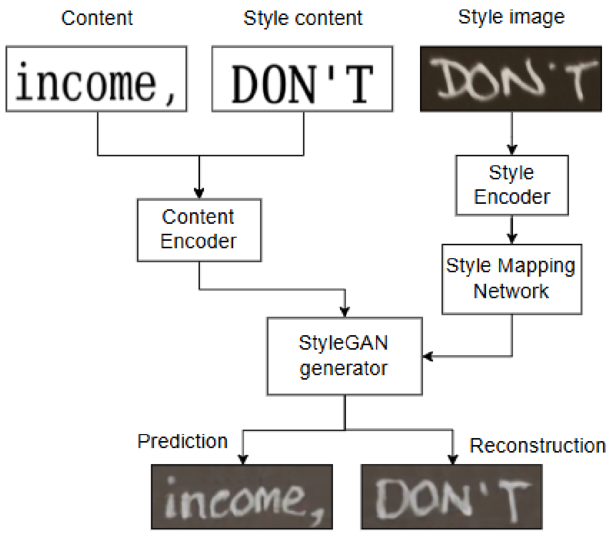


Fig. 5. High-level diagram of the model training process

1. Content — an image of the text we want to see on the generated image. This image is formed from the annotations to the images, namely, the text is taken and transformed into an image using a file that describes the font (VerilySerifMono) and functions from the PIL library. The size of all images is fixed — 192 by 64 pixels. This is done in order to have the same data types in all modules of our model, which greatly simplifies implementation.
2. Style content — an image of the text that appears on the image with the desired style.
3. Style image — an image that contains the desired style for the generated image.
4. Content Encoder — it is a pre-trained ResNet18 model that does not include average pooling layers and all subsequent ones, which is done to preserve the spatial properties of the image.
5. Style Encoder — it is a pre-trained ResNet18 model that lacks fully-connected layers.

6. Style mapping network — the key component of the StyleGAN model, which transforms the style vector obtained by the Style Encoder into a vector of parameters used to control the style on each layer of the generator using fully connected layers. The network structure consists of the following layers: PixelNorm (for normalizing the values of the vector obtained from the StyleEncoder component), Linear (fully connected layer), and LeakyReLU (nonlinear activation function).

7. StyleGAN generator — is the main component of the model that is responsible for generating images, taking as input an image with text and a vector of parameters obtained using the Style mapping network.

The synthesizing network consists of several convolutional units that gradually increase the image resolution starting from a small initial feature map. Each block uses stylized layers, such as AdaIN [10], to control the style of the image.

AdaIN (Adaptive Instance Normalization) — is a key component of StyleGAN that is responsible for styling images based on the style derived from the mapping network. AdaIN normalizes the feature activations for each feature map separately, and then scales and shifts their values according to the parameters obtained from the style vector.

Formally, the work of this layer can be represented as follows:

Let's say we have: x — input feature tensor for the convolutional layer; w — style vector obtained from the mapping network; $\mu(x)$ and $\sigma(x)$ — mean and standard deviation for any channel.

First, we calculate the affine transformation:

$$y_\sigma = A_\sigma w + b_\sigma, \;\; y_\mu = A_\mu \cdot + b_\mu$$

where A_σ and A_μ — weight matrices for multiplying the vector w, which provide different scaling depending on the style; b_σ and b_μ — displacement vectors that shift the parameters that control the mean value (for y_μ) and standard deviation (for y_σ) of features in the image space.

After the affine transformation, the style vector w creates the parameters y_σ and y_μ, which are then passed to AdaIN. Each feature channel in the tensor is normalized and modified according to these parameters, allowing each channel to respond to the style w.

Next, the features are normalized, where the mean $\mu(x)$ and standard deviation $\sigma(x)$ are calculated for each channel, and then the value is normalized:

$$x'_{ij} = \frac{x_{ij} - \mu(x)}{\sigma(x)}.$$

After normalization, each channel is scaled by y_σ and shifted by y_μ, which allows you to change the style:

$$x_{AdaIN} = y_\sigma x' + y_\mu$$

Thus, the affine transformation helps to transfer style parameters from the style vector w to the feature level, where they determine the color, texture, and other characteristics of the generated image. This allows for flexible control of visual attributes through the latent vector w.

8. Prediction — an image generated by the generator that combines a style from the Style image and text from the Content image.

9. Reconstruction — is an image generated by a generator that combines the style from the Style image and the text from the Style content image; under ideal conditions, the Reconstruction image should be identical to the Style image.

Next, we will describe additional components of the system that help calculate the loss function for our model.

Let's start with the discriminator (Fig. 6), as we can see, using the Discriminator component we calculate two loss functions — Discriminator adversarial loss and Generator adversarial loss. To calculate these two functions, we use the classical approach for GAN models, namely:

$$L_{discriminator} = \frac{1}{2}(MSE\,(D(z_{style\,image}),1) + MSE\,(D(G(z_{style\,content}, z_{style\,image})),0)),$$

where MSE — mean squared error; D — discriminator; G — generator; $z_{style\,image}$ and $z_{style\,content}$ — the results of the Style Encoder and Content Encoder components, respectively.

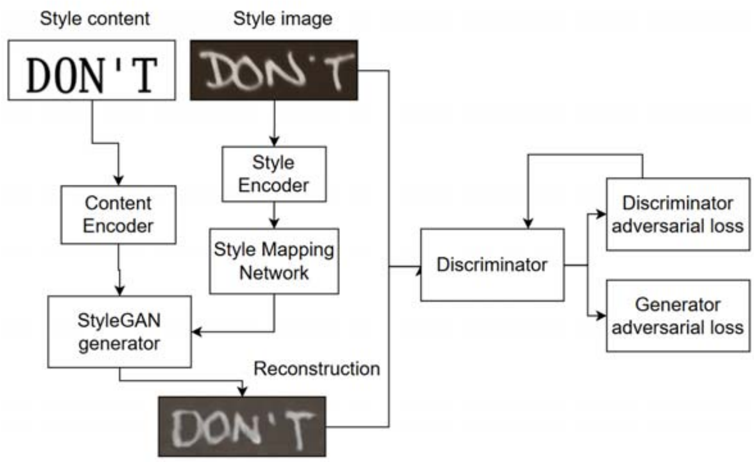


Fig. 6. Diagram of the process of calculating the competitive loss function for the discriminator and generator

The generator loss function is calculated as follows:

$$L_{generator} = MSE(D(G(z_{style\,content}, z_{style\,image})),1)).$$

To calculate the OCR loss function (Fig. 7), we use the trained TRBA (Text Recognition by Attention) model, which is denoted as OCR Recognizer in the diagram, Content label and Style content label are the text in the image Content and Style content, respectively, which are available during training.

In order to calculate the OCR loss function, we follow the following steps:

- Normalize the image size.
- Encode the text labels Content label and Style content label.
- Pass the Prediction and Reconstruction images to the model and get the predicted text on these images in the encoded form.
- Calculate the cross-entropy for the pairs (Prediction, Content label) and (Reconstruction, Style content label).
- Sum the obtained values and divide them by two.

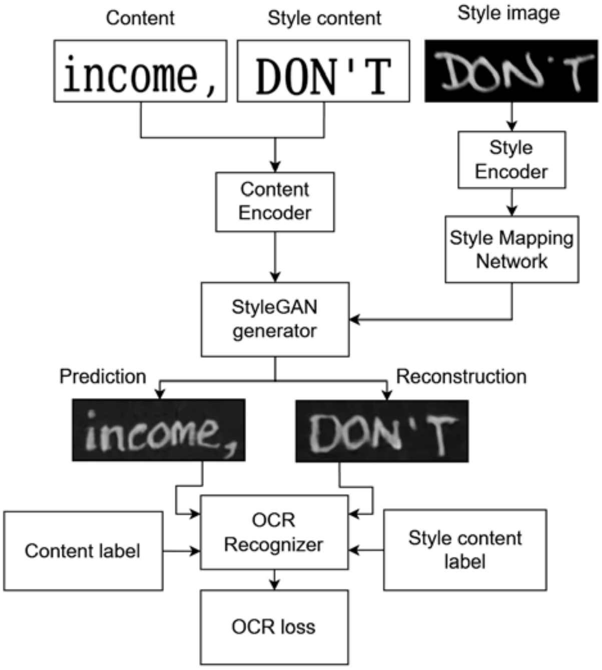


Fig. 7. Diagram of the OCR loss function calculation process

Formally, the process of calculating the OCR loss function can be represented as follows:

$$L_{OCR} = \frac{1}{2}(CE(TRBA(Prediction), Content\,label) +$$
$$+ CE()TRBA(Reconstruction), Style\,content\,label)\,,$$

where CE — cross entropy; $TRBA$ — pretrained model TRBA.

To calculate the Reconstruction of the loss function (Fig. 8), we need to calculate the difference between the Style image and the Reconstruction image. Formally, this is written as follows:

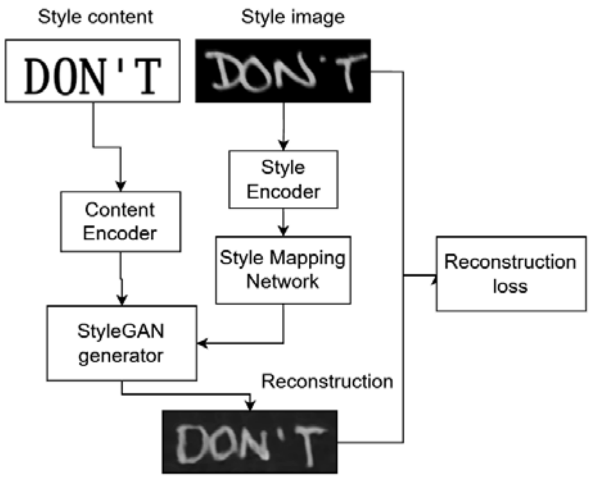


Fig. 8. Diagram of the process of calculating the Reconstruction loss function

$$L_{reconstruction} = L1\,(Style\,image, Reconstruction)\,,$$

where $L1$ — is the $L1$ loss function or mean absolute error.

To calculate the Cycle of the loss function (Fig. 9), we first need to get a Reconstruction image, and then pass it to the model input and get another image. The idea is that we should lose as little data as possible during the transformations driven by the model. Once we have two images, we calculate the average absolute error between them:

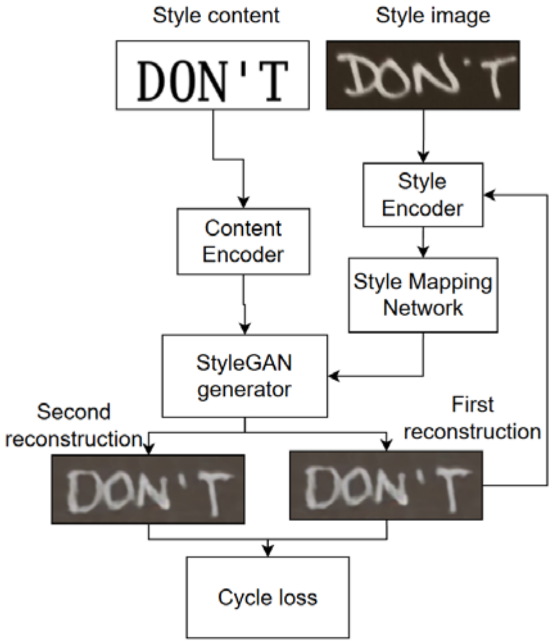


Fig. 9. Diagram of the process of calculating the Cycle of the loss function

To calculate the Typeface loss function (Fig. 10), we use the pre-trained VGG16 model, in which the last classification layer is removed, which means that the model will not perform classification, but will simply return a feature vector. After that, we pass two images to the model input — Style image and Reconstruction and calculate the L1 loss function between them, formally written as follows:

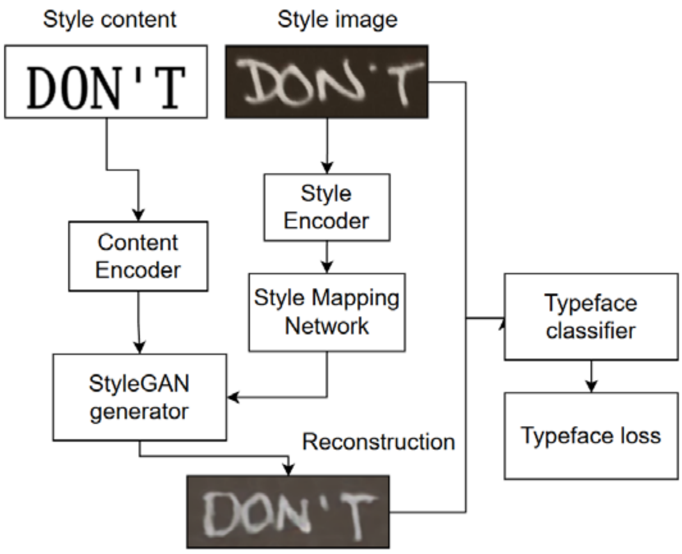


Fig. 10. Diagram of the Typeface loss function calculation process

$$L_{typeface} = L1(VGG16(Style\,image), VGG16(Reconstruction)).$$

To calculate the Perceptual and Texture loss functions (Fig. 11), we use the trained VGG16 model. For each block of VGG16, the difference between the activations of the corresponding layers of the input and target images is calculated, and the difference is measured using $L1$ loss, thus we obtain the Perceptual loss function:

$$L_{perceptual} = \frac{L1(X,Y)}{hw},$$

where X and Y — are activations at the i-th layer for the input and target images, respectively; h and w — height and width of the corresponding activations.

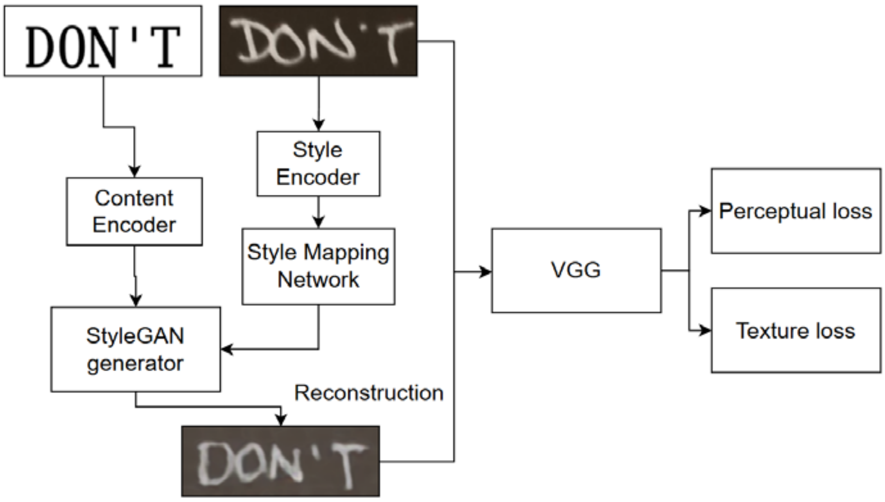


Fig. 11. Diagram of the process of calculating Perceptual and Texture loss functions

Texture losses are calculated in a similar way, but additionally, Gram matrices are calculated (The element G_{ij} represents the mutual correlation between the i-th and j-th channels. It calculates how similar these two channels are on average, i.e. how much the pixel value variations of one channel are related to the variations of the other. The gram matrix provides an idea of how different parts of the filters "cooperate" with each other to reflect the style or texture of the image):

$$L_{texture} = \frac{L1(G(X), G(Y))}{h\ w},$$

where $G(X)$ and $G(Y)$ — are Gram matrices, which are defined as $G(X) = X\,X^{\mathrm{T}}$.

The final loss function of the generator is as follows:

$$L_{final} = 0.06L_{generator} + 0.07L_{OCR} + 2.0L_{reconstruction} + 2.0L_{cycle} +$$

$$+ 0.0L_{typeface} + 25.0L_{perceptual} + 7.0L_{texture}.$$

The coefficients were selected empirically.

Verification and validation of the system engineering method of the system

The fourth stage involves verification and validation of the developed system to ensure that all technical specifications and stakeholder requirements are met. First of all, the model used to implement the Model Training component should be taken into account. It can be one of the following types of neural network architectures (this list is not exhaustive):

- convolutional neural networks;
- recurrent neural networks;
- generative-adversarial neural networks.

Next, you can analyze the functionality, adequacy, and performance of the developed system based on the selected benchmarking metric.

So, based on the definition of the concept of “method” [4], we can formulate the definition of the System Engineering Method of the conceptual model and the system of text replacement in images with style preservation: it is an ordered set of classes of tasks, processes, resources, business rules and relationships between them to produce a system based on the system engineering methodology, the Eriksson–Penker business profile and Data Science technologies.

Defining criteria for evaluating system performance

When designing a specific text-to-picture image replacement system, it is important to use appropriate metrics to evaluate the quality of image generation. The goal of such metrics is to measure how well the new images match the original in terms of visual similarity, stylistic authenticity, and detail accuracy. The metrics used in this paper are: mean square error (MSE), structural similarity (SSIM), peak signal-to-noise ratio (PSNR), and Frechet distance between distributions (FID). Each of them assesses different aspects of system quality and is effective for evaluating the performance of generative models.

1. Mean squared error (MSE)

Mean Squared Error (MSE) — is a metric that evaluates the difference between the pixel values in the original image and the generated image. It is calculated as the arithmetic mean of the square deviations between the corresponding pixels of the two images. The formula for calculating MSE looks like this:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(I_{\text{orig}}(i) - I_{\text{gen}}(i))^2 ,$$

where N — number of pixels in the image; $I_{\text{orig}}(i)$ and $I_{\text{gen}}(i)$ — pixel values of the original and generated images, respectively.

Advantages of MSE:

Easy to implement and interpret;

Well suited for evaluating pixel-by-pixel rendering accuracy.

Why it’s useful: MSE allows you to measure the exact difference between the original and the generated image, which is important for assessing the quality of preservation of text elements and image details after text replacement.

2. Structural similarity (SSIM)

Structural Similarity Index Measure (SSIM) — is a metric that evaluates the similarity between two images in terms of their structure, brightness, and contrast.

SSIM tries to model human perception, which makes it more sensitive to changes in image structure [11]. SSIM calculation formula:

$$SSIM(I_{\text{orig}}, I_{\text{gen}}) = \frac{(2\mu_{\text{orig}}\mu_{\text{gen}} + C_1)(2\sigma_{\text{orig-gen}} + C_2)}{(\mu_{\text{orig}}^2 + \mu_{\text{gen}}^2 + C_1)(\sigma_{\text{orig}}^2 + \sigma_{\text{gen}}^2 + C_2)},$$

where μ_{orig} and μ_{gen} — average pixel values of the original and generated images; σ_{orig} and σ_{gen} — covariance between the original and generated image; C_1 and C_2 — small constants to stabilize calculations.

Advantages of:

Good at showing structural changes in an image that are difficult to detect with MSE.

Simulates visual perception, making it more suitable for assessing image quality from the perspective of the human eye.

Why it's useful: SSIM helps to assess how well the basic structures and textures of the original image are preserved after text replacement, which is important for maintaining the style and authenticity of visual content.

3. Peak Signal-to-Noise Ratio (PSNR)

Peak Signal-to-Noise Ratio (PSNR) — is a metric used to evaluate the differences between an original and a generated image by measuring the level of noise or distortion. PSNR is calculated based on MSE and is expressed in decibels (dB) [12]:

$$\text{PSNR} = 10 \cdot \log_{10}\left(\frac{\text{MAX}^2}{\text{MSE}}\right),$$

where MAX — is the maximum possible pixel value in the image (for example, 255 for 8-bit images), and MSE is the root mean square error.

Advantages of PSNR:

Sensitive to small changes in brightness and colour.

A high PSNR value indicates that the image is close to the original.

Why it's useful: PSNR allows you to evaluate how well the system preserves brightness and contrast after text replacement, which is important for ensuring the visual quality of the image after editing.

4. Fréchet Inception Distance (FID)

Fréchet Inception Distance (FID) — is a metric used to evaluate the quality of generated images based on the similarity of distributions between real and generated images. FID measures the distance between the Gaussian distributions of features from images that are extracted using a pre-trained neural network (usually InceptionV3) [13]. FID calculation formula:

$$\text{FID} = \left\|\mu_{\text{orig}} - \mu_{\text{gen}}\right\|^2 + \text{Tr}\left(\Sigma_{\text{orig}} + \Sigma_{\text{gen}} - 2(\Sigma_{\text{orig}}\Sigma_{\text{gen}})^{1/2}\right),$$

where μ_{orig} and μ_{gen} — average feature values for original and generated images, Σ_{orig} and Σ_{gen} — covariation of these features.

Advantages of:

Evaluates not only pixel similarity, but also deeper features of images, making it more flexible for evaluating high-level stylistic characteristics.

Metric sensitivity to small changes that can be important for generating visually realistic images.

Why it's useful: FID is one of the main metrics for evaluating the quality of generative models such as StyleGAN. It allows you to evaluate how well the system is able to generate realistic images while preserving the visual and stylistic characteristics of the text and overall composition.

Conclusion. The use of MSE, SSIM, PSNR, and FID metrics provides a compreh ensive approach to evaluating the performance of a style-preserving text-to-picture system. MSE and PSNR evaluate the pixel-by-pixel reproduction accuracy, SSIM takes into account structural similarity, and FID allows you to determine how realistic the generated images look. Together, these metrics help to comprehensively evaluate the quality of the system and its ability to accurately reproduce the stylistic and visual features of images after text replacement.

Verification and validation of the system engineering method

The fourth stage involves verification and validation of the conceptual model and the developed version of the image text replacement system with style preservation. The model was built on the basis of the TextStyleBrush model architecture and trained for 50 epochs on a dataset consisting of 20 thousand segmented images. The system was also verified and validated, and the results are shown in the following Table 2.

Table 2. Comparative analysis of the use of metrics to assess the quality of the developed system

Metric	Test set 1	Test set 2	Test set 3	Average
MSE	0.0196	0.0229	0.0288	0.0238
SSIM	0.5071	0.5489	0.4460	0.5007
PSNR	15.3077	15.8087	15.4113	15.5092
FID	1.3876	0.7256	1.1510	1.0881

Interpreting the results of metrics such as MSE, SSIM, PSNR, and FID can give you an idea of the quality of the generated images compared to the original ones. Here is a breakdown of what each metric typically shows:

1. Mean squared error (MSE)

MSE quantifies the root mean square difference between the pixel values of two images. Lower values indicate better similarity. Although there is no absolute scale, an MSE value close to zero indicates high similarity, while higher values indicate greater divergence.

2. Structural similarity index (SSIM)

The SSIM ranges from -1 to 1, where 1 indicates complete similarity. An SSIM value of 0.5 indicates moderate structural similarity between images. In general, values above 0.7 are considered acceptable, and values above 0.9 are considered excellent.

3. Peak signal-to-noise ratio (PSNR)

PSNR is a logarithmic measure that compares the maximum possible signal power with the power of distorting noise. In image processing, a PSNR of 20–30 dB is generally considered acceptable, while a value of more than 30 dB is good and more than 40 dB is excellent. A PSNR value of 15 dB indicates poor quality, often indicating that the generated images are very different from the originals.

4. Fréchet Inception Distance (FID)

FID compares the distribution of generated images to real images. Lower FID scores indicate better quality and diversity of the generated images. Scores below 10 are generally considered good, while scores above 50 indicate poor quality. A FID score of 1.0881 indicates that your GAN is generating high quality images that are very similar to the real dataset.

Overall score:

- MSE: Low, which is good.
- SSIM: Moderate; can be improved.
- PSNR: Low; indicates noticeable differences in image quality.
- FID: Excellent; indicates good GAN performance.

Conclusion. MSE, SSIM, and PSNR metrics are used in almost every work related to generative networks, they are simple and quick to calculate, but such simplicity is usually not suitable for evaluating such systems, it is more done to be able to compare results with older works. It is better to focus on the results of the FID metric, which more accurately reflects the result for our particular problem.

CONCLUSIONS

1. A method of system engineering of a conceptual model and a system for replacing text in images with preservation of style is proposed, based on the modified Eriksson–Penker business profile [5–8] of the system's representation at the meta-level, as well as international standards of DataScience [2] and DataMining [3] processes, which is the basis for algorithmizing the development of specific system components. The effectiveness of the method is investigated on the example of developing a system for automatic text replacement in an image while preserving the style.

2. The proposed method of system engineering is aimed at developing specialized systems designed to replace text on design layouts and various products of companies and brands that need to adapt their visual materials for different markets and languages for promoting goods and services.

3. The use of the system engineering method significantly speeds up and streamlines the implementation of a particular system and reduces the cost of its development, verification and validation.

4. Prospects for further research are aimed at applying the system engineering method to implement a system based on other mathematical models, forming performance evaluation metrics and scientifically sound methods for verifying and validating the method.

REFERENCES

1. P.P. Maslianko, O.S. Maystrenko, "The system engineering of organizational system informatization projects," *KPI Sci. News*, no. 6, pp. 34–42, 2008.
2. C. O'Neil, R. Schutt, *Doing data science: Straight talk from the frontline*. O'Reilly Media, Inc., 2013, 406 p.
3. F. Provost, T. Fawcett, *Data science for business: What you need to know about data mining and data-analytic thinking*. O'Reilly Media, Inc., 2013.
4. Pavlo P. Maslianko, Yevhenii P. Sielskyi, "Method of system engineering of neural machine translation systems," *KPI Science News*, no. 2, pp. 46–55, 2021. doi: https://doi.org/10.20535/kpisn.2021.2.236939
5. H.-E. Eriksson, M. Penker, *Business modeling with UML*. New York: John Wiley & Sons, 2000, 459 p.

6. A. Kossiakoff, W. Sweet, S. Seymour, S. Biemer, *System Engineering Principles and Practice*. M.: DMK Press, 2014, 624 p.
7. D.K. Hitchins, *Systems Engineering: A 21st Century Systems Methodology*. Wiley, 2007, 528 p.
8. S.Krymskyi, "Metod," in *Filosofskyi Entsyklopedychnyi Slovnyk; V.I. Shynkaruk, Ed.* Kyiv, Ukraine: Abrys, 2002, 742 p. doi: https://doi.org/10.20535/ kpisn.2021.2.236939
9. Praveen Krishnan, Rama Kovvuri, Guan Pang, Boris Vassilev, Tal Hassner, "TextStyleBrush: Transfer of Text Aesthetics from a Single Example," *Journal of Latex Class Files*, vol. 14, no. 8, August 2015. Available: https://arxiv.org/pdf/ 2106.08385
10. X. Huang, S. Belongie, "Arbitrary Style Transfer in Real-Time with Adaptive Instance Normalization," *2017 IEEE International Conference on Computer Vision* (*ICCV*)*, Venice, Italy, 2017*, pp. 1510–1519. doi: https://doi.org/10.1109/ICCV.2017.167
11. Peter Ndajah, Hisakazu Kikuchi, Masahiro Yukawa, Hidenori Watanabe, Shogo Muramatsu, "SSIM image quality metric for denoised images," *International Conference on Visualization, Imaging and Simulation(VIS '10)*, pp. 53–57, 2010.
12. A. Horé, D. Ziou, "Image Quality Metrics: PSNR vs. SSIM," *20th International Conference on Pattern Recognition, Istanbul, Turkey, 2010*, pp. 2366–2369. doi: https://doi.org/10.1109/ICPR.2010.579
13. Yaniv Benny, Tomer Galanti, Sagie Benaim, Lior Wolf, "Evaluation Metrics for Conditional Image Generation," *International Journal of Computer Vision*, 129, pp. 1712–1731, 2021. doi: https://doi.org/10.1007/s11263-020-01424-w

Received 25.11.2024

INFORMATION ON THE ARTICLE

Pavlo P. Maslianko, ORCID: 0000-0003-4001-7811, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine, e-mail: masliankop@gmail.com

Mykola D. Romanov, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine, e-mail: kolya.romanov8@gmail.com

КОНЦЕПТУАЛЬНА МОДЕЛЬ ТА СИСТЕМА ДЛЯ ЗАМІНИ ТЕКСТУ НА ЗОБРАЖЕННІ ЗІ ЗБЕРЕЖЕННЯМ СТИЛЮ / П.П. Маслянко, М.Д. Романов

Анотація. Заміна тексту на зображенні, зокрема, зі збереженням його стилю, є складним завданням, яке потребує вирішення низки наукових задач та нових технічних рішень. Однією з основних проблем є збереження автентичності та гармонійності зображення після внесення змін. Мета дослідження — розроблення концептуальної моделі та системи заміни тексту на зображеннях зі збереженням його стилю на основі методології системної інженерії та бізнес-профілю Еріксона–Пенкера, забезпечуючи природну інтеграцію нових текстових елементів у контекст зображення. Методика реалізації — методологія системної інженерії і бізнес-профіль Еріксона–Пенкера для формалізації впорядкованого процесу розроблення системи для заміни тексту на зображенні зі збереженням стилю. Результати дослідження — метод розроблення системи на основі застосування технік системної інженерії, який складається з чотирьох основних етапів. На першому етапі структуру системи моделюють як бізнес-профіль Еріксона–Пенкера, на другому — визначають множину процесів, характерну для класу систем Data Science та міжнародного стандарту CRISP-DM, моделюють структурне і динамічне представлення концептуальної моделі системи та інтерфейси взаємодії компонентів, на третьому етапі виконують імплементацію конкретної версії системи, а на четвертому — верифікація та валідація системи. Запропоновано метод системної інженерії концептуальної моделі та системи заміни тексту на зображеннях зі збереженням його стилю, що ґрунтується на модифікованому бізнес-профілі Еріксона–Пенкера подання системи на метарівні, а також міжнародних стандартів процесів Data Science та Data Mining.

Ключові слова: метод системної інженерії, бізнес профіль Еріксона–Пенкера, концептуальна модель, система для заміни тексту на зображенні зі збереженням стилю.

UDC 519.688; 004.89; 004.9; 78.071.4: 78.078(477)+316.74
DOI: 10.20535/SRIT.2308-8893.2025.2.05

DETERMINING THE LEVEL OF PROPAGANDA IN OPERA LIBRETTOS USING DATA MINING AND MACHINE LEARNING

I. DATS, O. GAVRILENKO, K. FESHCHENKO

Abstract. The article presents an adapted multifactorial model that can be used to determine the level of propaganda in librettos to world operas. This model was created using the linear convolution method, for which eight indicators were selected that are most effective in identifying elements of propaganda in the text, taking into account the subject area's peculiarities. Each of the selected indicators was calculated using statistical analysis, data mining, and machine learning methods. As a result of applying the proposed method, the value function is calculated for each libretto, based on which a conclusion is made as to whether it contains elements of propaganda or not.

Keywords: art, propaganda, opera, libretto, multivariate model, statistical analysis, Data Mining, Machine Learning, information technology.

INTRODUCTION

Propaganda in art is the use of artistic forms to influence public opinion, shape ideas, and spread specific ideologies or political views. It can be both explicit and subtle, serving as an instrument of the state, religion, or social movements.

When studying the factors that influence human opinions in various areas of activity, it is worth paying attention to the vast and diverse realm of "agitation" in art. Since classical times, this has included visual and monumental art; during the Renaissance, masterpieces carried propaganda of a new era for humanity. Later, theatrical art acquired a dual meaning, while musical compositions and cinema, with their strong emotional impact, took on a special role in global propaganda.

The development of propaganda in art is based on:

- The promotion of an individual or a collective's creative activity (promotional advertising), which helped advance the careers of "useful" figures in the creative field.
- The involvement of specialists in the propaganda of artistic products, where musical content and literary foundations contributed to patriotic songwriting (particularly from the perspective of socialist state leaders).

Quite often, musical works used for propaganda incorporated compositions by other composers or folk songs, embedding entirely new meanings into them. For example:

The anthem of the USSR (at least its musical material) was taken from Mykola Lysenko's "Epic Fragment", whose impact and emotional depth made it highly suitable for Soviet state propaganda.

The agitational song "Far Beyond the River", which fully adopted a Ukrainian insurgent song about a fallen hero, was repurposed by the Red Army to promote the fight against what they considered old and bourgeois elements.

These are just a few examples of musical works that, in addition to raising issues of plagiarism in music, also highlight the problem of identifying propaganda.

Due to the vast diversity of art forms, this article focuses on the propagandistic impact on opera audiences, considering opera as a genre with a long history, an elite form of art, and a significant part of world culture.

Propaganda in opera has been particularly evident in productions staged in China [1], Nazi Germany, and the Soviet Union. For example, the works of Richard Wagner, which glorify ancient Germanic legends, were used to emphasize the superiority of the German nation and the Aryan race, reinforcing the ideology of world domination.

Similarly, the soviet regime implemented propaganda slogans by repurposing older russian operas and creating new, ideologically charged soviet works that praised and glorified the soviet government, its achievements, and way of life. In Ukraine: "The Death of the Squadron" by Yuliy Meitus, "Standard-Bearers" by Oleksandr Bilash. In russia: "In the Storm and Alpine Story" by Tikhon Khrennikov, as well as the film-operetta "Wedding in Malinovka" and the film-musical "Three Fat Men", among others.

Another intriguing aspect of musical propaganda is its presence in modern advertising. Commercials often feature simple, easily memorable melodies consisting of just a few notes, making them instantly recognizable and associated with the promoted product or message. In instrumental, vocal, and stage music, propaganda can be embedded in an emphasized form, calling for specific conclusions or even radical actions.

Overall, propaganda in opera has significant historical importance, particularly in societies where culture was used as a tool of ideological influence. As a synthesis of music, drama, and visual art, opera has a strong emotional impact, making it an effective medium for conveying political and ideological messages. Given the large volume of textual data in opera librettos and arias, identifying propagandistic elements requires advanced technologies.

Therefore, addressing this issue necessitates the integration of artistic expertise, including the work of playwrights, directors, actors, composers, and poets, along with information technologies such as mathematical modeling, Data Mining, statistical analysis, and Machine Learning techniques. This combination will enable systematic detection of propaganda in opera librettos, providing new insights into how ideological messages are embedded in classical and modern operatic works.

ANALYSIS OF LITERARY SOURCES AND PROBLEM STATEMENT

Research on propaganda detection demonstrates a variety of approaches and conclusions in this field. Scholars are increasingly leveraging modern techniques, particularly machine learning models such as *BERT* and *GPT*–4, to analyze and detect propaganda in textual data streams. These models can identify and classify different propaganda techniques across various texts.

Study [2] used a pre-trained *BERT* model to improve the detection of propaganda in news articles. The model processed text at the word level and integrated sentence-level features, effectively distinguishing between propagandistic and non-propagandistic content. However, issues such as data imbalance were identified, leading researchers to employ methods like oversampling and data augmentation to address them.

Study [3] focused on annotating and detecting propaganda using *GPT*–4. The research involved a multi-stage annotation process to ensure high-quality data, compiling a dataset of annotated paragraphs from diverse news sources to analyze propaganda techniques across different topics.

Study [4] examined the impact of propaganda on the political landscape in the U.S., revealing that disinformation in mass media significantly influenced social discourse and policymaking. This study proposed further research through ontology construction based on interdisciplinary methods from computer science and social sciences.

Study [5] conducted detailed text analysis, identifying 18 propaganda techniques in manually annotated news articles. The research also introduced a new BERT-based neural network to enhance propaganda detection.

Study [6] presented a credibility assessment methodology for questionable information, using semantic similarity metrics on knowledge graphs to calculate the shortest paths between conceptual nodes.

Study [7] explored the history and evolution of information warfare methodologies, comparing American, British, and Russian models while introducing the concept of "semantic warfare" in the modern world.

A crucial limitation of current machine learning models is their reliance on supervised learning, meaning they require human-labeled training datasets. This introduces an element of subjectivity, as the classification of certain texts as propaganda depends on human judgment.

Additionally, social media plays a significant role in propaganda dissemination today [8]. For example:

Study [9] introduced the *CatRevenge* model, designed to identify active and passive revenge communication in social media, which aligns with propaganda detection. The model used *Slangzy* (an internet slang dictionary) for preprocessing, assigning *TF–IDF*-based weights to words and employing a *CATBoost* classifier to reduce overfitting.

Study [10] investigated influential individuals in knowledge-sharing processes within internal social networks, predicting future knowledge flow patterns and analyzing propaganda's ideological impact through a four-phase methodology combining social network analysis and structural modeling.

Study [11] analyzed how social media posts by influential figures affected cryptocurrency markets, highlighting an example of propaganda in commerce.

Study [12] deals with the problem of detecting propaganda in text files. The authors consider methods for solving the problem of classifying textual information for spam filtering, contextual advertising, news categorization, and creating thematic catalogs.

Study [13] presents a multifactorial model for determining the level of propaganda in a publication. The publications used were text news and social media posts. The model was created based on the linear convolution method. This model considered 10 indicators, a high level of each of which indicates the presence of propaganda in the publication. This model is based only on statistical data and calculations made using Data Mining, statistical analysis and Decision Theory algorithms.

Study [14] provides an overview of multilingual models for working with limited data sets and analyzes their development. The following models are considered: *XLM–RoBERTa*, *mBERT*, *LASER*, *MUSE*.

These studies emphasize the importance of using sophisticated Machine Learning, Statistical Analysis, Data Mining, and careful data annotation processes to detect and analyze propaganda. They provide valuable insights into methodologies that can improve the accuracy and reliability of propaganda detection systems, which is crucial for understanding and mitigating the impact of propaganda.

It should be noted that the process of propaganda detection continues to require the development of various mathematical models to better identify this form of communication.

In addition, it should be emphasized that none of the proposed models has been used to identify propaganda in the musical and theatrical arts in general and opera in particular.

The authors of this study propose a modified version of the Multi-Factor Propaganda Detection Model (MMDP) from Study [13], adapted specifically for evaluating propaganda levels in opera librettos. Additionally, the study examines how propaganda detection results from MMDP correlate with the assessments of opera experts. An information technology was developed to conduct experimental research. By integrating Machine Learning, Statistical Analysis, and Data Mining with artistic expertise, this study aims to fill the existing gap in identifying propaganda in opera as an elite and historically significant art form.

OBJECTIVE AND TASKS OF THE RESEARCH

The objective of this research is to adapt the MMDP [13] for processing and analyzing the libretto of world operas to identify signs of propaganda within them. To achieve this objective, the following tasks have been set:

- Compile a dataset of libretto from well-known world operas that differs from the dataset presented in study [13].
- Select from the 10 propaganda indicators outlined in study [13] those that are most relevant to the chosen artistic domain.
- Improve methods for determining indicators that are characteristic of propaganda detection in publications.
- Utilize the MMDP to calculate the level of propaganda content within the compiled dataset.
- Draw conclusions regarding the presence of propaganda indicators in the libretto texts.

MATERIALS AND METHODS OF RESEARCH

The object of the study is the process of identifying propaganda in opera libretto (hereafter referred to as publications) based on an analysis of information about them. Specifically, the study considers the following factors:

- Primary source of the publication (in this context, the literary work that served as the basis for the opera libretto).
- Brief description of the primary source.
- Word count in the publication.
- Sentence count in the publication.
- Syllable count in the publication.

- Total number of opera productions currently available on streaming platforms.
- Number of productions of operas based on libretto contained in the dataset.
- Number of reviews of opera performances based on the libretto in the dataset.
- Number of re-posts of the publication (in this context, the number of video recordings of the opera based on the given libretto on a streaming platform).
- Number of likes under the video recording of the opera based on the given libretto on a streaming platform.
- Number of comments under the video recording of the opera based on the given libretto on a streaming platform.
- Sources of re-posts (in this context, channels that share opera video recordings based on the given libretto).

The set of publications and all necessary information for this research was obtained from [15].

The successful completion of the study requires both basic statistical data and data obtained using Data Mining and Machine Learning techniques

Fig. 1 illustrates the main steps involved in determining the level of propaganda in publications.

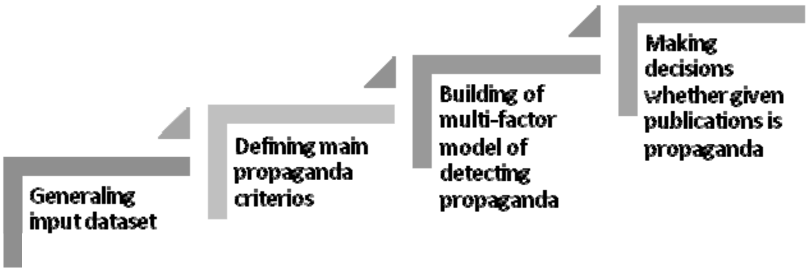


Fig. 1. Main Steps for Determining the Level of Propaganda in Publications

The proposed propaganda detection principle is based on calculating a metric that reflects the degree of correspondence between a given publication and pre-selected propaganda indicators. This is achieved using the convolution method.

To compute the values of the indicators, the study employs statistical analysis methods, as well as Data Mining and Machine Learning techniques. Additionally, specialized software was developed for conducting intelligent analysis and obtaining results based on these methods.

ADAPTATION OF A MULTIFACTOR MODEL FOR CALCULATING THE LEVEL OF PROPAGANDA IN OPERA LIBRETTO

The process of constructing a multifactor model for calculating the level of propaganda in publications, based on the convolution method, can be outlined in the following stages [16; 17].

Stage 0: Preprocessing of the publication text.

Stage 1: Calculation of numerical indicators for the model.

Stage 2: Calculation of importance coefficients for each indicator.

Stage 3: Calculation of the value function.

Stage 4: Formulation of conclusions regarding whether the given publication is propagandistic.

These stages are illustrated in Fig. 2.

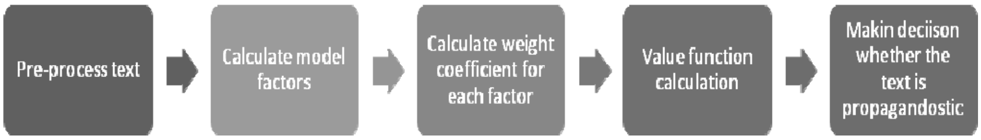


Fig. 2. The process of building a multifactor model

Step 0

Input: a set of publications. $P=(P_1,\ldots,P_l)$.

Output: sets of words $A_1, A_2,\ldots,A_l$ used in the publications $P_1,\ldots,P_l$ respectively.

To form each set, it is recommended to preprocess the text of the publications using lemmatization and stemming processes. This helps reduce the size of the set by eliminating root-related words and auxiliary parts of speech.

PHASE 1

Input: a set of publications $P=(P_1,\ldots,P_l)$ a set of propaganda features $X=(x_1,\ldots,x_8)$ [13]:

x_1 = {attempts to manipulate the audience};

x_2 = {the publication is aimed at evoking emotions};

x_3 = {frequent repetition of a specific idea in the publication};

x_4 = {frequent reposting of the publication};

x_5 = {simplicity of the publication's text};

x_6 = {a high level of propaganda in the original source};

x_7 = {belonging to a specific topic that is particularly susceptible to propaganda};

x_8 = {the publication has an impact on the viewer}.

It is necessary to calculate the levels of propaganda for each of the given features.

At the output, a set is formed $K^j=(K_1^j,\ldots,K_8^j)$, where K_i^j, $i=1,\ldots,8;$ $j=1,\ldots,l$ — the values of the metrics that indicate the level of propaganda in publication according to feature x_i.

1. **Calculation of the metric** K_1^j. A numerical assessment of manipulative attempts in texts can be based on methods of computational linguistics, sentiment analysis, lexical analysis, and machine learning.

Emotional tone analysis. Manipulative texts often contain emotionally charged words (e.g., fear, threats, exaltation). Emotional dictionaries $(SentiWordNet, LIWC, NRC, VADER)$ are widely used to determine the emotional tone of a text.

For example, if a publication contains a negative tone (fear, anger), manipulation is possible. If a publication contains excessive positivity, propaganda is possible. If the negative emotion index is higher, and the aggregated sentiment score is too low, deliberate escalation is possible.

Detection of logical fallacies and manipulative techniques. Manipulators use certain rhetorical techniques:

- Appeal to Fear (e.g., the phrase “Either you are mine, or death!” from G. Puccini’s opera Tosca).
- False Dilemma (e.g., the phrase “Who does not fall at my feet will perish!” from G. Verdi’s opera Nabucco).
- Ad Hominem (e.g., the phrase “God, who has placed a ray of His divinity within us, created man to rule!” from G. Verdi’s opera Don Carlos).

Lexical patterns and Machine Learning methods are used to detect emotional fallacies. The model is trained on datasets containing labeled manipulative phrases. If a text contains excessively negative predictions, it may be an attempt at manipulation.

Lexical Analysis: Frequency of Manipulative Constructions. Manipulative texts often contain:

- Generalizations (“Everyone knows this!” from G. Verdi’s opera Rigoletto; “No sinner will escape God’s judgment!” from G. Verdi’s opera Don Carlos).
- Evaluative Judgments (“There has never been a more ruthless tyrant!” from G. Verdi’s opera The Sicilian Vespers).
- Appeals to Authority (“The law is the law!” from G. Puccini’s opera Tosca).

If a text contains many generalizations and emotionally charged evaluative judgments, it may be manipulative.

Text Style Analysis (Stylometry). Manipulative texts may contain a high number of exclamations, many interrogative sentences (rhetorical questions), as well as excessively long or very short sentences.

Thus, score K_1^j for a publication P_j, $j = 1,\ldots,l$ is calculated as follows:

$$K_1^j = \alpha_1 S^j + \beta_1 L^j + \gamma_1 F^j + \delta_1 C^j, \quad (1)$$

where S^j — sentiment of the text, determined using a word dictionary with specific polarity (positive, negative, neutral) ($S^j = 1$ for a positive or negative tone, $S^j = 0$ for neutral text); L^j — relative frequency of manipulative clichés (lexical features) compared to their total variety; F^j — relative frequency of logical fallacies (fallacies detection) compared to their total variety; C^j – relative frequency of identified stylistic characteristics (stylometry) compared to their total variety; $\alpha_1, \beta_1, \gamma_1, \delta_1$ — weight coefficients. In this study $\alpha_1 = 0{,}4$; $\beta_1 = 0{,}3$; $\gamma_1 = 0{,}2$; $\delta_1 = 0{,}1$. The values of the weight coefficients, as well as those in the subsequent models, were chosen according to the specifics of the subject area and agreed upon with an expert — M.I. Hamkalo, director of a musical-dramatic theater and associate professor at the Tchaikovsky National Music Academy of Ukraine.

It is evident that $0 \le K_1^j \le 1$, and the closer its value is to one, the more manipulative features the given publication contains. Thus, based on x_1 criterion, it can be considered propagandistic.

2. **Calculation of the Metric** K_2^j. The emotional orientation of a text indicates the extent to which it evokes specific emotions (fear, joy, anger, etc.). It can be assessed using the following approaches:

1. Sentiment Analysis.

2. Emotion Detection.
3. Lexical Analysis of Emotional Intensity.
4. Deep Learning (NLP -models) (models: $BERT, GPT, LSTM$).

In this study, sentiment analysis was used to evaluate the emotional orientation of the text.

Thus, the metric K_2^j for a publication P_j, $j = 1,\ldots,l$ is calculated as the overall emotional score:

$$K_2^j = \alpha_2 S^j + \beta_2 E^j + \gamma_2 C^j, \quad (2)$$

where S^j — sentiment of the text (this parameter was described earlier); E^j — proportion of emotional words in the text; C^j — atext style analysis (this parameter was also described earlier); $\alpha_2, \beta_2, \gamma_2$ — weight coefficients. In this study $\alpha_2 = 0{,}5$; $\beta_2 = 0{,}3$; $\gamma_2 = 0{,}2$.

It is evident that $0 \le K_2^j \le 1$ and the closer its value is to one, the higher the level of emotional intensity in the given publication. Thus, based on this criterion, it can be considered propagandistic.

3. **Calculation of the Metric** K_3^j. If an idea is expressed using different words, vector models $(Word2Vec, BERT)$ can be used to find similar expressions.

In this study, a vector model $Word2Vec$ was used, with cosine similarity as the similarity measure. Thus, the metric K_3^j for a publication P_j, $j = 1,\ldots,l$ is calculated as follows:

$$K_3^j = \cos(\theta) = \frac{(B^j D^j)}{\| B^j \| \| D^j \|}, \quad (3)$$

where B^j and D^j — are vectors representing objects (word vectors extracted from the publication j); $(B^j D^j)$ — is the dot product of the vectors; $\| B^j \|$, $\| D^j \|$ — are the magnitudes (norms) of the vectors; $\cos(\theta)$ — represents the cosine of the angle between the vectors.

It is evident that $0 \le K_3^j \le 1$ and the closer its value is to one, the more frequently a particular idea is repeated in the given publication. Thus, based on this criterion, it can be considered propagandistic.

4. **Calculation of the Metric** K_4^j**.** The frequency of reposting a publication refers to the number of video recordings of opera performances based on the analyzed libretto found on streaming platforms (Netflix, YouTube, etc.).

Thus, the metric K_4^j for a publication P_j, $j = 1,\ldots,l$ is calculated as the relative frequency of the opera performance's P_j on a streaming platform using the following formula [18; 19]:

$$K_4^j = \frac{n_j}{n}, \quad (4)$$

where n_j — the number of video recordings of the opera based on the given libretto j; n — the total number of operas found on the platform.

It is evident that $0 \le K_4^j \le 1$ and the closer its value is to **one**, the more frequently the given publication is reposted. Thus, based on this criterion, it can be considered propagandistic.

It should be noted that the accuracy of this metric K_4^j depends on the choice of the streaming platform. The more popular the platform, the larger audience it covers within the study. On the other hand, major platforms require processing a large volume of statistical data, which may introduce additional complexities in calculating this metric.

For example, on the OperaVision website [20], 264 video recordings of opera performances were found. G. Verdi's opera Aida was represented in 8 videos. Thus, for the libretto of this opera, $K_4^j \approx 0{,}03$. On other platforms, this metric may have a different value due to variations in statistical data.

5. **Calculation of the Metric** K_5^j. This metric indicates the readability of the given publication's text.

The metric K_5^j for a publication P_j, $j = 1,\ldots,l$ is calculated as follows:

$$K_5^j = \left(206{,}835 - 1{,}015\frac{a_j}{b_j} - 84{,}6\frac{c_j}{a_j}\right)0{,}01, \qquad (5)$$

де a_j — total number of words; b_j — total number of sentences; c_j — total number of syllables.

This metric K_5^j is known as the Flesch Reading Ease Index [21].

The interpretation of this metric's values is shown in Table 1.

Table 1. Interpretation of Flesch Reading Ease Index Values

Score	School level	Notes
0,9 – 1,0	Grade 5	Very easy to read. Easily understood by an average 11-year-old student
0,9 – 0,8	Grade 6	Easy to read. Conversational language for consumers
0,8 – 0,7	Grade 7	Fairly easy to read
0,7 – 0,6	Grades 8-9	Standard language. Easily understood by 13–15-year-old students
0,6 – 0,5	Grades 10-12	Fairly difficult to read
0,5 – 0,3	College	Difficult to read
0,3 – 0,1	Technical Graduate	Very difficult to read. Best understood by university graduates
0,1 – 0,0	Professional	Extremely difficult to read. Best understood by university graduates

It is evident that $0 \le K_5^j \le 1$ and the closer its value is to one, the easier the given publication is to read. Thus, based on this criterion, it can be recommended as propagandistic.

6. **Calculation of the Metric** K_6^j. The primary source refers to the literary work that served as the basis for the libretto (publication).

The metric K_6^j for a publication P_j, $j = 1,\ldots,l$ is calculated as follows:

Step 1. Identify the primary source.

Step 2. Find a brief description of this work.

Step 3. Use a model $Word2Vec$ to determine the key words from the text description.

Step 4. Calculate the cosine similarity (equation 3) between the key word vector and a predefined reference vector.

$$Q = (fight; danger; trait; enemy; alliance; tragedy;$$
$$destruction; patriot; glory; unite; fame;$$
$$recretion; hero; unbeatable) . \quad (6)$$

In equation (6), the vector Q used in this study was constructed based on a set of words characteristic of propaganda detection. It was reviewed and approved by an expert — M.I. Hamkalo, director of a musical-dramatic theater and associate professor at the Tchaikovsky National Music Academy of Ukraine. This vector can be adjusted or modified depending on the specific subject area of analysis.

It is evident that $0 \le K_6^j \le 1$, and the closer its value is to one, the higher the likelihood that the given publication has a propagandistic nature. Thus, based on criterion x_6, it can be considered propagandistic.

As an example, we can consider the opera "The Golden Ring" by Ukrainian composer Borys Lyatoshynsky, based on the libretto by Yakiv Mamontiv, which was inspired by Ivan Franko's novel "Zakhar Berkut". It is well known that the novel contains a call to struggle against external and internal enemies. This leitmotif was transferred into the libretto and, consequently, into the opera.

Thus, according to criterion x_6, the opera "The Golden Ring" exhibits propaganda elements.

7. **Calculation of the Metric** K_7^j. Consider a publication $P_j, j = 1,\ldots,l$; the set of words used in the publication A_j; the set of topics $S = (s_1; s_2; \ldots; s_r)$,in which propagandistic publications are most frequently found, and the dictionaries of characteristic words for these topics $T_1; T_2; \ldots; T_r$. The topics and their corresponding dictionaries should be predefined. Some of these topics include:

- Politics: "power", "tyranny", "monarchy", "autocracy", "rebellion", "discord", "revolutionary movement", "coup", "betrayal", "intrigue", "enemies", "opponents",...
- Military Conflicts: "army", "legion", "foreign rule", "tyranny of conquerors", "conquest",...
- Ideology: "people", "nation", "society", "unity", "solidarity", "cohesion", "alliance", "threat", "danger", "monarchy",...
- Conspiracies and Disinformation: "the real truth", "triumphant truth", "secret conspiracy", "treacherous plan", "spies", "accomplices",...

The metric K_7^j indicates whether the publication P_j belongs to one of the topics in the set S. It is calculated as follows:

Step 1. Compute the Jaccard similarity coefficients between the set of words A_j and each topic dictionary T_k, $k = 1, 2, \ldots, r$ [22]:

$$J(A_j, T_k) = \frac{|A_j \cap T_k|}{|A_j \cup T_k|}; \quad (7)$$

Step 2. Select the maximum Jaccard coefficient:

$$J_{max}(A_j, T_k) = \max_{k=1,2,\ldots,r} J(A_j, T_k), \quad j = 1,\ldots,l$$

and establish which topic T corresponds to this maximum value $s_k \in S$.

Step 3. The metric K_7^j is defined as:

$$K_7^j = J_{\max}(A_j, T_k). \quad (8)$$

It is evident that $0 \le K_7^j \le 1$ and the closer its value is to one, the more closely the publication aligns with topics that are most susceptible to propaganda. Thus, based on this criterion x_7 it can be considered propagandistic.

As an example, we can again consider the opera "The Golden Ring". The libretto of this opera can be categorized under the "Ideology" topic, which is frequently influenced by propaganda. Therefore, for this libretto (publication), the value of the metric K_7^j is quite close to 1.

8. **Calculation of the Metric** K_8^j. To assess the audience reach and its impact, the overall score is calculated as follows:

$$K_8^j = \lambda_1 X_1 + \lambda_2 X_2 + \lambda_3 X_3, \quad (9)$$

where X_1 — relative number of likes to the total number of opera views; X_2 — proportion of opera views relative to the most popular opera in the dataset; X_3 — relative number of comments to the total number of opera views; $\lambda_1, \lambda_2, \lambda_3$ — weight coefficients.

It is evident that $0 \le K_8^j \le 1$ and the closer its value is to one, the greater the level of influence the given publication has on the audience. Thus, based on criterion x_8, it can be considered propagandistic.

It should be noted that the accuracy of the metric K_8^j similar to K_4^j depends on the choice of the streaming platform.

PHASE 2

Input: indicators K_i^j, $i = 1,\ldots,8$; $j = 1,\ldots,l$, calculated using formulas (1)–(9).

It is necessary to calculate importance coefficients for each criterion to determine the value function.

Output: coefficients ω_i.

To compute these coefficients ω_i, the following steps must be performed:

Step 1: Form statistical samples from the indicators K_i^j with corresponding names.

Step 2: Select a threshold value, exceeding which a publication can be considered propagandistic.

In this study, by analogy with Chaddock's scale [18; 19], which defines the strength of correlation between two random variables, the following scaling was proposed:

0,0 – 0,1 — no propaganda;

0,1 – 0,3 — low level of propaganda;

0,3 – 0,5 — noticeable level of propaganda;

0,5 – 0,7 — moderate level of propaganda;

0,7 – 0,9 — high level of propaganda;

0,9 – 1,0 — very high level of propaganda.

In this study, all levels of propaganda starting from the noticeable level were considered. Thus, the threshold value was set at $\overline{K_i} = 0{,}3$.

This threshold was introduced to facilitate further statistical calculations and ensure the convenient comparison of results with expert opinions.

It should be noted that no universally defined percentage threshold exists in scientific sources that explicitly determines when a text is considered propagandistic [23]. This study emphasizes the importance of qualitative analysis and the recognition of specific influence techniques rather than establishing a universal quantitative threshold.

In future research, a more personalized approach is planned for each propaganda characteristic.

Step 3: If $K_i^j \geq \overline{K_i}$, the given publication P_j is considered propagandistic based on the feature x_i. Otherwise, it is classified as non-propagandistic. Each publication is assigned the value «1», if it is propaganda based on this feature, and «0» otherwise.

$$P_j \to \widetilde{K}_i^j = \begin{cases} 1, & \text{if } K_i^j \geq \overline{K_i}; \\ 0, & \text{if } K_i^j < \overline{K_i}. \end{cases}$$

The transition from quantitative values K_i^j to boolean functions $\widetilde{K}_i^j$ was made to facilitate the comparison of results with expert opinions.

Step 4: Calculate the Relative Frequency of Propagandistic Publications for Each Feature x_i.

$$w_i = \frac{m_i}{n},$$

where m_i — the number of propagandistic publications based on feature x_i; n — the total number of publications in the dataset.

Step 5: Normalize the Relative Frequencies w_i:

$$\omega_i = \frac{w_i}{w_1 + w_2 + \ldots + w_8}.$$

PHASE 3

Input: a set of publications $P = (P_1, \ldots, P_l)$, indicators K_i^j, $i = 1, \ldots, 8$; $j = 1, \ldots, l$ and coefficients ω_i.

It is necessary to calculate the value function for each publication to determine the presence of propaganda features.

Output: the value function result V_j .

The value function V_j , s computed using the linear aggregation method as follows [16; 17]:

$$V_j = \sum_{i=1}^{8} (\omega_i K_i^j). \tag{10}$$

Based on the values of V_j a **statistical sample** of value function results is formed according to equation (10):

$$V = (V_1, V_{2,}, \ldots, V_l).$$

PHASE 4

Input: a set of publications $P = (P_1, \ldots, P_l)$ and a statistical sample $V = (V_1, V_{2,}, \ldots, V_l)$ (see Step 3).

Output: conclusions regarding which publications $P = (P_1, \ldots, P_l)$ єare propagandistic.

Recommendations are made according to the following rule [24]:

- If $V_j \geq \overline{V}$, $(j = 1, \ldots, l)$, then the publication P_j is recommended as propagandistic.
- If $V_j < \overline{V}$, $(j = 1, \ldots, l)$, then the publication P_j is not recommended as propagandistic.

In this rule $\overline{V} = 0{,}3$ — is the threshold value for the sample V (analogous to Step 2).

$$P_j \to \widetilde{K}_i^j = \begin{cases} 1, & \text{if } K_i^j \geq \overline{K_i}; \\ 0, & \text{if } K_i^j < \overline{K_i}. \end{cases}$$

Thus, a publication is assigned «1», if it is considered propaganda and «0» otherwise.

The correctness of the provided conclusions is evaluated using the *Recall* та *Precision* metrics:

$$Precision = \frac{tp}{tp + fp}, \quad Recall = \frac{tp}{tp + fn},$$

where tp — the number of correctly identified propagandistic publications (true positives); fp — the number of incorrectly identified propagandistic publications (false positives); fn — the number of incorrectly identified non-propagandistic publications (false negatives).

OBTAINED RESULTS

As part of this study, a dataset was compiled, containing the librettos of 10 operas (Table 2).

For these operas, the value function was calculated, based on which conclusions were drawn regarding the presence of propaganda elements in their librettos.

Table 2. Compiled Dataset

Libretto	Opera Title	Composer
P_1	The Huguenots	Giacomo Meyerbeer
P_2	The Mastersingers of Nuremberg	Richard Wagner
P_3	Fidelio	Ludwig van Beethoven
P_4	The Troubadour	Giuseppe Verdi
P_5	A Life for the Tsarc	Mikhail Glinka
P_6	La Traviata	Giuseppe Verdi
P_7	Carmen	Georges Bizet
P_8	Madame Butterfly	Giacomo Puccini
P_9	Turandot	Giacomo Puccini
P_{10}	The Marriage of Figaro	Wolfgang Amadeus Mozart

The obtained results are presented in Table 3.

Table 3. Obtained Results

Libretto	K_1^j	K_2^j	K_3^j	K_4^j	K_5^j	K_6^j	K_7^j	K_8^j	V_j
P_1	1	1	1	0	1	0	1	1	1
P_2	1	1	1	0	1	0	1	1	1
P_3	1	1	1	0	1	0	1	1	1
P_4	1	1	1	1	1	1	1	1	1
P_5	1	1	1	1	1	1	1	1	1
P_6	0	1	0	1	0	0	0	0	0
P_7	0	1	0	1	0	0	0	0	0
P_8	0	1	0	1	0	0	0	0	0
P_9	0	1	0	1	0	0	0	0	0
P_{10}	0	1	0	1	1	0	0	0	0

In Table 3, each publication P_j, $j = 1,\ldots,10$ is assigned a value «1», if it is considered propaganda based on feature x_i, $i = 1,\ldots,8$ and value «0» otherwise.

DISCUSSION OF RESEARCH RESULTS

The obtained results were compared with the expert opinion of M.I. Hamkalo, associate professor in the field of musical directing at the Tchaikovsky National Music Academy of Ukraine. The comparison is presented in Table 4.

Thus, from Table 4, it is evident that the proposed MMDP identified the presence of propaganda elements in the same opera librettos as the expert. Accordingly, the values of the $Precision = 1$, $Recall = 1$ metrics, confirm the high accuracy of the MMDP.

T a b l e 4 . Comparison of MMDP Results with Expert Opinion

Libretto	V_j	Expert Opinion	Expert's Argumentation
P_1	1	1	Propaganda: Anti-Catholicism
P_2	1	1	Propaganda: German nationalism
P_3	1	1	Propaganda: Liberalism and the struggle for freedom
P_4	1	1	Propaganda: Revolutionary spirit and fight for independence
P_5	1	1	Propaganda: Russian imperial narrative
P_6	0	0	No propaganda: Pure melodrama about personal emotions, without political or social context
P_7	0	0	No propaganda: The opera has no ideological connotations, only depicting emotions and the fatality of destiny
P_8	0	0	No propaganda: A personal tragedy and cultural misunderstandings, without a political message
P_9	0	0	No propaganda: A mythical story not tied to specific political events
P_{10}	0	0	No propaganda: Despite criticism of the feudal system, it is more about romantic twists than politics

CONCLUSIONS

Propaganda in opera is a powerful tool for influencing society, utilizing the impact of music, librettos, and stage performances to shape specific ideological narratives. Throughout different historical periods, opera has served as an instrument of state propaganda, expressing political, social, and nationalist ideas.

In the XIX century, during the era of Romanticism, opera was often used to elevate national spirit and support struggles for independence (for example, "Nabucco" by Giuseppe Verdi became a symbol of the Italian liberation movement). In the XX century, totalitarian regimes actively employed opera to reinforce state ideology: Soviet socialist realism, Nazi Germany, and Maoist China promoted productions that glorified the party, leaders, or the "ideal citizen".

Despite this, opera also served as a means of protest and counter-propaganda. It became a tool for criticizing authority or social structures, often using allegorical plots or hidden messages.

Thus, opera not only reflects historical context but also actively shapes public consciousness, making it a significant instrument of both official and oppositional propaganda.

This study presents an adapted multifactor model, which allows for the assessment of propaganda levels in the librettos of world opera masterpieces. This model is based on the linear aggregation method, for the implementation of which eight indicators were selected. These indicators are the most effective in detecting propaganda elements in a text, taking into account the specific features of the subject area. Each of the selected indicators was calculated using statistical analysis, Data Mining methods, and Machine Learning techniques. As a result of the proposed method, a value function is computed for each publication, based on which a conclusion is drawn regarding whether it contains propaganda elements or not.

Advantages of the Proposed Model:

1. Elimination of Human (Subjective) Influence — the model's calculations rely solely on statistical data or data obtained through Data Mining and Machine Learning methods, ensuring objectivity in detecting propaganda indicators.

2. Scalability — the model can be easily expanded by adding new indicators or removing outdated ones, making it adaptable to evolving research needs.

3. Result Accuracy — the correctness of the obtained results is guaranteed by the use of classical Data Mining and Machine Learning methods.

Disadvantages of the Proposed Model:

1. Large Data Requirements — the model requires the collection and storage of vast amounts of statistical and textual data, which may pose challenges in data management.

2. Continuous Accuracy Monitoring — the reliability of conclusions must be regularly evaluated. In this study, an expert in the subject area was consulted. In other domains, the accuracy of the MMDP model should be validated using multiple propaganda detection methods.

The obtained results can be used as an effective tool in information warfare, both in Ukraine and globally, serving as a powerful element of intent analysis. Additionally, they can assist directors and actors in musical-dramatic theaters, including opera houses and operetta theaters.

Focusing specifically on the concept of artistic propaganda, the proposed methodology can be applied to all forms of art that are in some way related to textual data, such as songs, films, theater, literature, and poetry. For these domains, the methodology would differ only in terms of input statistical data, such as song lyrics, brief descriptions of literary works, or play scripts. It would also vary in the values of weight coefficients in formulas (1), (2), and (9), as well as in the adaptation of propaganda features presented in [13], where some characteristics may be added or removed depending on the specific artistic field.

REFERENCES

1. Zongrui Zhang, "Model Opera" of the 20th Century in Chinese Musical Culture," *Art History Notes*, no. 43, pp. 206–210, 2023. doi: https://doi.org/10.32461/2226-2180.43.2023.286862
2. W. Li, S. Li, C. Liu, L. Lu, Z. Shi, S. Wen, "Span identification and technique classification of propaganda in news articles," *Complex Intell. Syst.*, vol. 8, pp. 3603–3612, 2022. doi: https://doi.org/10.1007/s40747-021-00393-y
3. Maram Hasanain, Fatema Ahmed, Firoj Alam, "Can GPT-4 Identify Propaganda? Annotation and Detection of Propaganda Spans in News Articles," *Computation and Language (cs.CL)*, 2024. doi: https://doi.org/10.48550/arXiv.2402.17478
4. K. Hamilton, "Towards an Ontology for Propaganda Detection in News Articles," in *R. Verborgh et al. The Semantic Web: ESWC 2021 Satellite Events. ESWC 2021. Lecture Notes in Computer Science*, vol. 12739. Springer, Cham, 2021. doi: https://doi.org/10.1007/978-3-030-80418-3_35
5. G. Da San Martino, S. Yu, A. Barrón-Cedeño, R. Petrov, P. Nakov, "Fine-grained analysis of propaganda in news article," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, Association for Computational Linguistics, pp. 5635–5645, 2019. doi: https://doi.org/10.18653/ v1/D19-1565
6. G.L. Ciampaglia, P. Shiralkar, L.M. Rocha, J. Bollen, F. Menczer, A. Flammini, "Computational fact checking from knowledge networks," *PLoS One*, 10(6), 15, 2015. doi: https://doi.org/10.1371/journal.pone.0128193
7. G. Pocheptsov, *Modern information wars*. Kyiv: Kyiv-Mogylianska Academy, 2015, 497 p.
8. S. Ghosal, A. Jain, "CatRevenge: towards effective revenge text detection in online social media with paragraph embedding and CATBoost," *Multimed Tools Appl.*, 83, pp. 89607–89633, 2024. doi: https://doi.org/10.1007/s11042-024-18791-y
9. R. Alhajj, J. Rokne (Eds), *Encyclopedia of Social Network Analysis and Mining*. Springer, New York, NY., 2018, 2200 p. doi: https://doi.org/10.1007/978-1-4614-7163-9
10. Ramona-Diana Leon, Raúl Rodríguez-Rodríguez, Pedro Gómez-Gasquet, Josefa Mula, "Social network analysis: A tool for evaluating and predicting future knowledge flows from an insurance organization," *Technological Forecasting and Social Change*, vol. 114, pp. 103–118, 2017. doi: https://doi.org/10.1016/j.techfore.2016.07.032

11. Sergii Telenyk, Grzegorz Nowakowski, Olena Gavrilenko, Mykhailo Miahkyi, Olena Khalus, "Analysis of the influence of posts of famous people in social networks on the cryptocurrency course," *Bulletin of the Polish Academy of Sciences Technical Sciences*, vol. 72(4), 2024. doi: https://doi.org/10.24425/bpasts.2024.150117
12. O. Gavrilenko, Y. Oliinyk, H. Khanko, "Analysis of Propaganda Elements Detecting Algorithms in Text Data," in *Z. Hu, S. Petoukhov, I. Dychka, M. He, (Eds) Advances in Computer Science for Engineering and Education II. ICCSEEA 2019. Advances in Intelligent Systems and Computing*, vol. 938, Springer, Cham, 2020, pp. 438–447. doi: https://doi.org/10.1007/978-3-030-16621-2_41
13. O. Gavrilenko, K. Feshchenko, "Detecting propaganda in news flows," *Adaptive systems of automatic control*, no. 1 (46), pp. 160–177, 2025. doi: https://doi.org/10.20535/1560-8956.46.2025.323759
14. V. Oliinyk, I. Matviichuk, "Low-resource text classification using cross-lingual models for bullying detection in the Ukrainian language," *Adaptive systems of automatic control*, no. 1 (42), pp. 87–100, 2023. doi: https://doi.org/10.20535/1560-8956.42.2023.279093
15. Opera librettos and arias by foreign authors notes. Accessed on: Feb. 23, 2025. [Online]. Available: https://musicinukrainian.wordpress.com/biblio/import_opera/
16. Jürgen Branke, Kalyanmoy Deb, Kaisa Miettinen, Roman Słowiński, *Multiobjective Optimization*. Springer-Verlag Berlin Heidelberg, 2008, 470 p. doi: https://doi.org/10.1007/978-3-540-88908-3
17. Kalyanmoy Deb, *Multi-Objective Optimization using Evolutionary Algorithms*. Wiley, 2001, 536 p.
18. Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, E.Ye. Keying, *Probability and Statistics for Engineers and Scientists*; 9th ed. Pearson, 2016, 816 p.
19. Sheldon Ross, *A First Course in Probability*; 10th ed. Pearson, 2018, 528 p.
20. *Operavision*. Accessed on: Feb. 23, 2025. [Online]. Available: https://operavision.eu/
21. Rudolf Flesch, *How to Write Plain English: A Book for Lawyers and Consumers*. Harper & Row, 1979, 126 p.
22. Jure Leskovec, Anand Rajaraman, Jeffrey D. Ullman, *Mining of Massive Datasets*. Cambridge University Press, 2014, 326 p.
23. Giovanni Da San Martino, Seunghak Yu, Alberto Barrón-Cedeño, Rostislav Petrov, Preslav Nakov, "Fine-Grained Analysis of Propaganda in News Articles," *Computation and Language* (*cs.CL*), 2019. doi: https://doi.org/10.48550/arXiv.1910.02517
24. P.G. Preethi, V. Uma, A. Kumar, "Temporal Sentiment Analysis and Causal Rules Extraction from Tweets for Event Prediction," *Procedia Computer Science*, no. 48, pp. 84–89, 2015. doi: https://doi.org/10.1016/j.procs.2015.04.154

Received 01.03.2025

INFORMATION OF THE ARTICLE

Iryna V. Dats, ORCID: 0000-0003-3851-2047, Tchaikovsky National Music Academy of Ukraine, Ukraine, e-mail: irynadats@gmail.com

Olena V. Gavrilenko, ORCID: 0000-0003-0413-6274, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine, e-mail: gelena1980@gmail.com

Kyrylo Yu. Feshchenko, ORCID: 0009-0002-8142-179X, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine, e-mail: fkirill440@gmail.com

ВИЗНАЧЕННЯ РІВНЯ ПРОПАГАНДИ В ОПЕРНИХ ЛІБРЕТО ЗА ДОПОМОГОЮ ЗАСОБІВ DATA MINING ТА MACHINE LEARNING / І.В. Даць, О.В. Гавриленко, К.Ю. Фещенко

Анотація. Подано адаптовану багатофакторну модель, яку можна використати для визначення рівня пропаганди в лібрето до світових опер. Модель створено на основі методу лінійної згортки, для реалізації якого обрано 8 індикаторів, найбільш ефективних для виявлення елементів пропаганди в тексті з урахуванням особливостей предметної галузі. Кожного з обраних індикаторів розраховано з використанням методів статистичного аналізу, Data Mining та машинного навчання. У результаті застосування запропонованого методу для кожного лібрето розраховується значення функції цінності, на основі якого робиться висновок про те, чи містить вона елементи пропаганди, чи ні.

Ключові слова: мистецтво, пропаганда, опера, лібрето, багатофакторна модель, статистичний аналіз, Data Mining, Machine Learning, інформаційна технологія.

UDC 623.7, 519.2
DOI: 10.20535/SRIT.2308-8893.2025.2.06

RANKING OF THE TECHNICAL CONDITION OF AIRCRAFT ACCORDING TO THE DIAGNOSTIC DATA OF THE GLIDER DESIGN

A. BOLOHIN, Yu. BOLOHINA, Yu. TYMCHUK

Abstract. Analysis of the data obtained through non-destructive testing of the power elements of the airframe structure enables the division of the studied aircraft into separate groups based on operational methods and, accordingly, predicts the probability of trouble-free operation of the airframe structure. Based on a bionic logical analysis of data on the technical condition of the aircraft's airframe design, the article considers methodological approaches to building the author's mathematical model of ranking the aircraft fleet. The ranking of the technical condition of aircraft serves as the basis for solving the problem of preparation, making a decision on extending the service life and taking into account the aging of the airframe design in operating conditions. The application of the proposed set of methods will create and establish a practical and effective system of intelligent support for decision-making on aircraft operations.

Keywords: technical condition of the airframe, strength, operation of aircraft, control systems, decision-making.

INTRODUCTION

Digital systems with classical methods of management, which are mainly focused on practice, except for the frequent productivity of micro-processing systems. A significant increase in the efficiency of such systems can be reached by the implementation of adaptive control algorithms. Adaptation of a living organization, which is designed to change the internal and external minds, is an analogue for the construction of the structures of technical systems with the establishment of the system of control for the change of internal and external parameters up to the criterion of regulation. Based on the classical methods of management, for the implementation of an adaptive control system, the necessary visibility of the great need for the analysis of the preliminary information for the skin-specific regulator, to develop a mathematical model and an adequate algorithm for assessing the parameters of the model, which is to add a folding problem for traditional approaches.

Fuzzy control is based on the rules formed by an expert who has some experience in a particular technical direction, and from the point of view of control system technology, fuzzy control is a regulation with a nonlinear gear response.

Thus, a management system based on fuzzy logic is not blind control using one or another law of regulation, but management with a current analysis of the situation and meaningful decision-making to issue a regulation signal, guided by common sense, that is, the most adequate response to a comprehensively assessed control error. These are minimal transients. This is high flexibility, reliability and efficiency and simplification of control compared to adaptive and self-learning control algorithms [1].

TASK STATEMENT

Management of the aircraft operation process requires the fulfillment of the conditions for ensuring the required level of flight safety and involves the adoption of organizational and technical decisions to maintain aircraft in good condition. The resource and reliability of aircraft is largely determined by the technical condition of the carrier elements (CE) of the airframe design. The results of the CE assessment according to the control data allow distributing aircraft into separate groups (classes), for which there are flight safety solutions obtained by performing strength calculations and based on operational experience. In practice, the known methods and algorithms for making decisions on the further operation of aircraft are reduced to the problem of formal choice of an alternative [2].

When making decisions, three main tasks are solved:

- ordering alternatives;
- distribution of alternatives by classes of solutions;
- highlighting a better alternative.

For example, such classes and, accordingly, solutions can be:

- aircraft meet flight safety requirements and do not require additional restoration work;
- it is necessary to carry out restoration work of various purposes and scope (repair at aircraft repair enterprises, extension of the assigned service life, transfer to operation according to technical condition, etc.);
- further operation of the aircraft is inexpedient due to the loss of the necessary operational characteristics of the CE of the airframe design.

The imperfection of the known methods of decision-making lies in the fact that information about the technical condition of the aircraft is characterized by the peculiarity of incompleteness and limited volumes of operational data, which significantly affects the validity of decisions in the management of the aircraft operation process.

Problematic issues of classification of aircraft into separate groups depending on the results of airframe control carried out in the conditions of operating parts have not been studied in technical diagnostics. Therefore, there is a need to develop methods for classifying aircraft that will increase the reliability of decisions made by an official (decision-maker – hereinafter referred to as DM) for a specific situation based on the results of control inspections of the airframe.

One of the directions for the development of methods for classifying airframe is the development of a mathematical model of bionic processing of experimental data, which simulates the mechanisms of functioning of biological systems and allows DM to understand the validity and expediency of various options for decisions when making decisions.

ANALYSIS OF RECENT RESEARCH AND PUBLICATIONS

Decision support systems are a qualitatively new level of automation of management processes in various areas of human activity.

Thus, in the work [3], a classification model has been developed that is able to accurately identify diseases and provide appropriate recommendations for patients. The proposed system uses a multi-level classification mechanism with appropriate prediction of changes in the patient's vital parameters.

In the process of operation of aviation equipment, as a result of control inspections, many values of parameters are determined that characterize and allow to assess the strength of the aircraft airframe structure.

In geometric interpretation, the technical condition of the airframe can be represented as a point in multidimensional space, the coordinate axes of which correspond to the control parameters.

The mathematical model of such a description of the technical condition of the airframe does not allow to provide the DM with the necessary data on the regularities of the airframe operation process in an accessible visual form, since it does not take into account the peculiarities of the mechanisms of human information processing.

The paper [4] shows that a person perceives the environment with the help of an internal figurative representation, his imaginary model is aimed at analyzing information by generalizing, identifying the dependence between its individual parts, searching for associative connections with previously made decisions.

The internal model of DM is difficult to identify and formally describe the transformations that are carried out when generalizing the acting factors and forming images, transforming multidimensional data and obtaining estimates of objects in a one-dimensional scale of values (from "minimum" to "maximum", from "bad" to "good", etc.).

In the same work, it is noted that when processing control data with statistical methods of pattern recognition, intuitive and understandable judgments of people about the trends in changing the parameters of the processes under study are used, that is, there is a transition to one-dimensional scales. In the works on psychology, the position that reveals the features of human processing of information in the form of images is substantiated. This pattern allows you to develop the structure of the bionic model, consisting of an interconnected set of modules that take into account subjective aspects and allow you to simulate the processing of information in the semantic image space of the subject area of the environment.

PRESENTATION OF THE MAIN MATERIAL

When solving the problems of monitoring the technical condition of the CE of the airframe structure, the subject area consists of numerical and linguistic values of the database on crack sizes, depth and area of corrosion, dimensions of deformation of the CE, coordinates of the damage site, normative data on tolerances, service life, etc.

The use of identification procedures in the assessment of technical condition allows you to evaluate parameters when their dimensionality is precisely known. At the same time, the calculation of the difference between the reference and the actual state, i.e. the performance of a comparison, is a necessary operation of the process of localization of the technical state. In addition, the number of parameters that characterize the technical condition is often unknown. Therefore, the disadvantages of existing diagnostic technologies and decision-making systems include the need to form requirements for experimental information about dynamic changes in the system, the use of which will allow classification and comparison of possible technical states.

The structure of the bionic model, consisting of separate modules, shown in Figure 1, uses the principles of vector information processing by a universal mathematical model for biological systems, as well as the principles of logical transformations of information performed by a person during the solution of intellectual problems [5].

To solve the problems of classification of aircraft according to the data of airframe parameter control, three vector spaces F_1, F_2, F_3 are needed, the dimensionality of which is determined by the number of parameters of the subject area of control.

The first space F_1 is created by estimates of the indicators of loss of strength of the CE based on the data of visual and non-destructive testing means of each individual CE, the results of computational and experimental studies of the strength of the CE (Fig. 1).

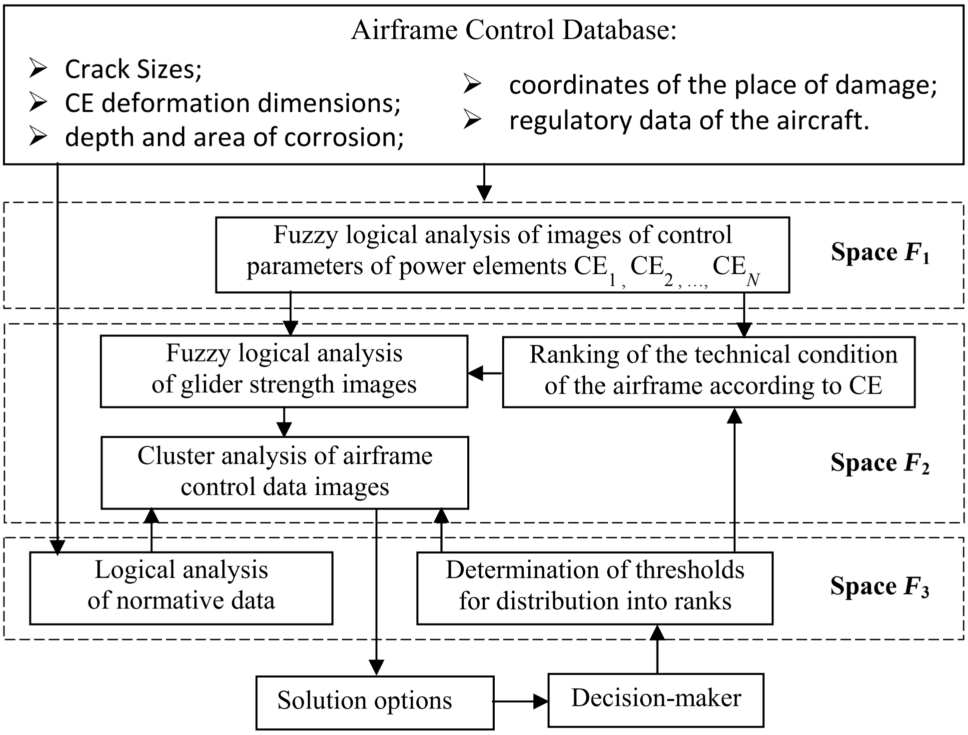


Fig. 1. Structure of the bionic model of data analysis

Due to the incompleteness and limited volume of operational data, the results of assessments of CE strength loss indicators are characterized by the presence of intervals of uncertainty of values. In addition, the assessments contain errors due to the subjective experience of a person working with control materials. Therefore, final decisions on strength loss assessments require amendments and adjustments.

Such clarifications can be carried out taking into account the use of psychological qualities of a person, use the criterion of approximate similarity of objects, i.e. vague logical conclusions when searching for such objects. The work [6] shows that the property of fuzziness of logical thinking helps a person to solve practical problems.

In the modules of the bionic model, which solve the problem of estimating the indicators of loss of strength of CE, operators for processing fuzzy logical dependencies based on the methods of fuzzy set theory are used [7]. Performing a fuzzy analysis of the control results allows you to transform the multidimensional description of the control data and move to a one-dimensional rating scale.

The model provides for the use of the second vector space F_2, where, based on the results of the control of individual CE, a generally fuzzy logical analysis of data images of the technical condition of the entire airframe structure is performed.

Determination of cases in which the values of the strength factors go beyond the tolerances is carried out using the methods of mathematical binary logic, which uses logical operators for calculating predicates.

In the space F_2 the property of a person to group and rank CE according to control data is simulated. The dimension of the space F_2 is determined by the number N – the total number of CE.

The number of CE classes that are formed depends on the values of the thresholds set by the DM. Therefore, the corresponding module of the bionic model also uses the method of fuzzy information processing, which takes into account the subjective representations of the DM.

The CE classification module uses an algorithm based on the principles of taxonomy of objects described in the paper [7] and takes into account that a person perceives the results of classification regardless of the geometric shape of the class in the multidimensional space of features.

Space F_3 takes into account additional information about the airframe operation process, it is here that the ranking of the technical condition and the preparation of draft decisions are carried out. The dimension of the space F_3 is determined by the number of normative numerical data and estimates of linguistic variables obtained by calculation when controlling the loss of strength.

Next, the construction of a mathematical model of the ranking of the technical condition based on a fuzzy logical analysis of the parameters of the glider control is considered.

In the module of fuzzy logical analysis of images of control parameters CE, a linguistic variable is introduced d_f = “difference in strength CE from the norm”. The universal set of d_f is the segment [0 1], and the set of values for d_f is the term set $G = \{G_1, G_2, G_3, G_4, G_5\}$, where G_1 — practically coincide; G_2 — small difference; G_3 — average difference; G_4 — big difference; G_5 — the maximum difference.

The number of discrete states of the term set G is chosen in accordance with the ability of the human operator to simultaneously perform operations with 5... 7 objects that belong to different images.

As a function of belonging $\mu(d_f)$, the formula (1) of a trapezoidal fuzzy number $d_f = (a_1, a_2, a_3, a_4)$:

$$\mu(d_f) = \begin{cases} 0, & \text{if} \quad d_f < a_1; \\ \mu(d_f) = \exp(-\exp(-(d_f - C_1))), & \text{if} \quad a_1 \le d_f < a_2; \\ 1, & \text{if} \quad a_2 \le d_f \le a_3; \\ \mu(d_f) = \exp(-\exp(-(d_f - C_2))), & \text{if} \quad a_3 < d_f \le a_4; \\ 0, & \text{if} \quad d_f > a_4; \end{cases} \quad (1)$$

where C_1, C_2 — the displacement coefficients of the function (1) along the axis d_f.

The sides of the trapezoid are formed on the basis of the functional dependence

$$\mu(d_f) = \exp(-\exp(-d_f)). \quad (2)$$

The dependence (2) approximates the assessment of observations in various subject areas of the experimenters' real decisions on the admissibility of research results. Dependence (2) has the beneficial properties of continuity, monotony, and smoothness.

The value of the belonging function $\mu(d_f)$ is perceived as a measure of the truth of the term G_i, which is established by experts in the subject area of aircraft operation. Fig. 2 shows the distribution of belonging functions to the linguistic variable d_f.

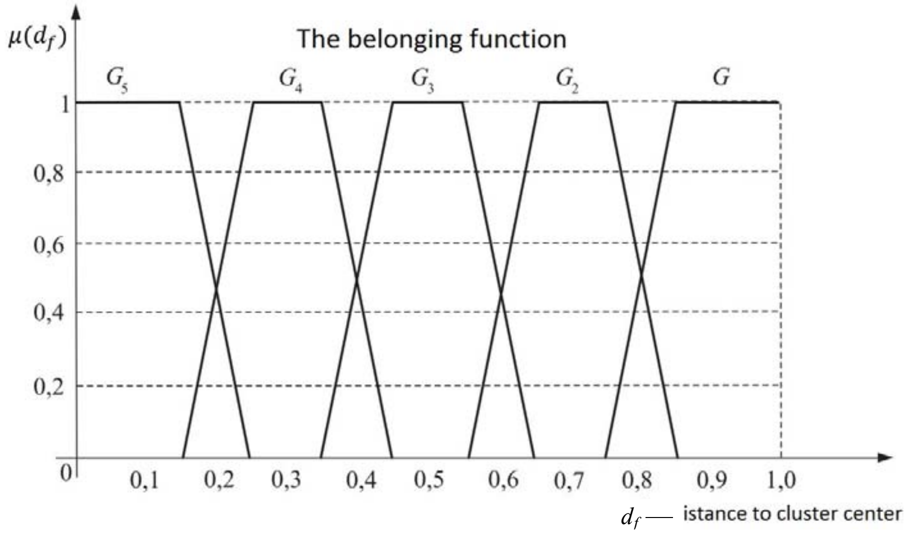


Fig. 2. Membership functions of subsets of the term set G

To analyze the dependence of the distance d_f on the parameters of the v_i of the first vector space F_1, a linguistic variable B_i is introduced, which, by analogy with the term set G, consists of the following terms:

B_{i1} — very small difference between the parameters of v_i;

B_{i2} — a slight difference between the parameters of v_i;

B_{i3} — average level of difference in v_i parameters;

B_{i4} — a significant difference between v_i parameters;

B_{i5} — a very significant difference between v_i parameters.

The volume of the database used during processing is determined by the elements of the set $V = \{v_1, v_2, ..., v_l\}$, where the list of parameters v_i $(i = 1,...,l)$. The number of parameters l is set by experts taking into account informal methods of describing the subject area.

It is also assumed that each linguistic variable is described by a trapezoidal fuzzy number $v_i = (a_1, a_2, a_3, a_4)$. Thus, a two-dimensional array of fuzzy linguistic variables $B_{i\kappa}$ $(i = 1,...,l;\ \kappa = 1,...,5)$ is created.

To form the rules for the transition from the values of the parameters to the values of the variable itself d_f in accordance with [7], an approximation of fuzzy logical inference rules is carried out using the formula

$$d_f = \sum_{k=1}^{5} h_k \overline{d_{f_\kappa}} \,,$$

where $h_k = \sum_{i=1}^{l} (1/l) \mu_{lk}$ characterizes the influence of each term of the linguistic variable G_k; $\overline{d_{f\,k}}$ — is the middle of the interval on the d_f axis, at which the values of the term $G_k \in (a_{k1}, a_{k4})$ are non-zero.

The module of fuzzy logical analysis of glider strength images is built by the method of designing and debugging fuzzy knowledge bases, which are sets of linguistic or logical operators.

The logical formula for determining the value of the loss of strength σ the airflame is built on the basis of the logical operator "TA" $(\wedge)$ and thevalues of the term set G of individual CE_i ($i = 1, ..., N$):

$$\sigma = \min \mathop{\wedge}_{i=1}^{N} d_{fi} \,.$$

The module for ranking the technical condition of gliders by CE selects such homogeneous subsets in the initial multidimensional data so that the objects within the groups are similar in known content to each other and have the same dimension, and the objects from different groups are not similar.

The module uses an algorithm that provides for the construction of the shortest open path between the vertices of the graph in the space F_2. The distance between the vertices of the graph is calculated by the method of potential functions [8].

Mathematical models of F_3 space modules are developed using methods that are used in F_2 space modules.

Effective application of the proposed mathematical model in the existing operation system can be, for example, when determining the timing of periodic maintenance. Thus, depending on the actual technical condition and the results of ranking by power elements, the deadline for work may be postponed. For example, for a specific aircraft, the technical condition of which is characterized by such variables as G_1, B_{i1}, i.e. loss of CE strength, there is no deterioration of the condition, the design is ranked as serviceable, the airframe meets the requirements of flight safety and does not require additional restoration work, the DM in the technical condition management system decided to increase the terms of performance of the relevant periodic work from 3 months to 6. This made it possible to save and efficiently allocate the resources necessary for the technical support of the operation process.

CONCLUSION

The application of the developed mathematical model for classifying the ranking of the technical condition on the basis of bionic logical analysis of airframe control data as part of the mathematical support information systems for managing the technical condition of aircraft will improve the quality of organizational and technical decisions made.

REFERENCES

1. A.A. Sherba, W.W. Kirick, *Systems with fuzzy logic for regulating electrical modes.* K.: Institute of Electrodynamics of the National Academy of Sciences of Ukraine, 2011.
2. B.M. Herasimov, W.M. Lokazyuck, O.G. Ohsiyuck, O.W. Pomorova, *Intelligent decision support systems*. K.: Publishing House of the European University, 2007.
3. Vedna Sharma, Surender Singh Samant, “A multi-level decision-making frame-work for heart-related disease prediction and recommendation,” *System Research and Information Technologies*, no. 4, 2023. doi: https://doi.org/10.20535/SRIT.2308-8893.2023.4.01
4. E.Yu. Barzylovych, M.V. Savenkov, *Statistical methods of assessing the state of aviation technology*. M.: Transport, 1987.
5. Yu.P. Shabanov-Kushnarenko, *Theory of intelligence: Mathematical tools*. Kharkiv: Higher school. Publishing House at Kharkiv University, 1984.
6. *Fuzzy sets in models of control and artificial intelligence*; Edited by D.A. Pospelova. M.: Nauka, 1986.
7. L.K. Konysheva, D.M. Nazarov, *Fundamentals of the theory of fuzzy sets: Tutorial*. St. Petersburg: Piter, 2011.
8. M.A. Aizerman, E.M. Braverman, L.I. Rozonoer, *The method of potential functions in the theory of machine learning*. M.: Science, 1970.

Received 05.04.2024

INFORMATION ON THE ARTICLE

Andrii S. Bolohin, ORCID: 0000-0003-0656-7799, State Research Institute of Aviation, Ukraine, e-mail: bolas@ukr.net

Yuliya O. Bolohina, ORCID: 0009-0004-7251-7137, State Research Institute of Aviation, Ukraine, e-mail: jbologina@ukr.net

Yurii M. Tymchuk, ORCID: 0009-0002-0425-1985, 4993rd Military Representation of the Ministry of Defense of Ukraine, Ukraine, e-mail: timab-d1972@ukr.net

РАНЖУВАННЯ ТЕХНІЧНОГО СТАНУ ПОВІТРЯНИХ СУДЕН ЗА ДАНИМИ ДІАГНОСТУВАННЯ КОНСТРУКЦІЇ ПЛАНЕРА / А.С. Бологін, Ю.О. Бологіна, Ю.М. Тимчук

Анотація. Аналіз даних, які отримують за допомогою засобів неруйнівного контролю силових елементів конструкції планера, дозволяє розподілити досліджені повітряні судна на окремі групи залежно від методів експлуатації і, відповідно, прогнозувати ймовірність безвідмовної роботи подальшої експлуатації конструкції планера. На основі біонічного логічного аналізу даних про технічний стан конструкції планера літака розглянуто методичні підходи побудови авторської математичної моделі ранжування парку повітряних суден. Ранжування технічного стану повітряних суден покладено в основу вирішення завдання підготовки прийняття рішення про продовження терміну служби з урахуванням старіння конструкції планера в умовах експлуатації. Застосування запропонованого комплексу методів дозволить створити і налагодити дієву та ефективну систему інтелектуального підтримання прийняття рішень про експлуатацію повітряних суден.

Ключові слова: технічний стан планера, міцність, експлуатація літальних апаратів, системи керування, прийняття рішень.

UDC 519.24
DOI: 10.20535/SRIT.2308-8893.2025.2.07

TIME SERIES FORECASTING USING THE NORMALIZATION MODEL

VIKTOR BONDARENKO, VALERIIA BONDARENKO

Abstract. Empirical constructions of time series models based on the reduction of initial data to normally distributed values have been proposed. The goal of a normalization method is to construct an optimal forecast that is linear for the updated data, and the forecasted original data is recovered through the inverse transformation. The different variants of such transformations have been considered, including the reduction of initial data to Gaussian fractional Brownian motion and a one-dimensional transformation using a strictly monotonic function. The computational experiment based on real data, which allows for a stationary model, confirms the higher quality of the forecast by the normalization method compared to traditional models.

Keywords: optimal forecast, stochastic model, parameter estimation, fractional Brownian motion.

PRELIMINARY INFORMATION AND STATEMENT OF THE PROBLEM

Time forecasting is defined as estimation the future values of some function of a time variable based on known observations up to the current moment. Other-words, if we observe the trajectory $x(t)$, $0 \le t \le T$, then it is necessary to evaluate the value $x(s)$, $T \le s \le T+\tau$. The estimated values are called forecasts and are denoted by $\hat{x}(s)$, . The stochasticity of the trajectory is an essential circumstance, so $x(\cdot)$ does not represent a deterministic function.

As a rule, the trajectory values are observed at discrete moments

$$0 \le t_1 < t_2 < \ldots < t_{n-1} < t_n \le T,$$

that is, the terms of the sequence are known $\{x_1, x_2, \ldots, x_n\}$, where $x_k = x(t_k)$, and "future" values are to be assessed $x(t_{n+1}), \ldots, (t_{n+r})$,

$$T \le t_{n+1} < t_{n+2} < \ldots < t_{n+r} \le T + \tau .$$

If it is known a priori that $x(t)$ is an implementation of a random process $\xi(t)$, which corresponds to a finite-dimensional distribution

$$P\{\xi(t_1) < x_1, \ldots, \xi(t_m) < x_m\} = \int_{-\infty}^{x_1} dy_1 \int_{-\infty}^{x_2} dy_2 \ldots \int_{-\infty}^{x_m} f_m(t_1,..,t_m, y_1, y_2,..y_m)\, dy_m$$

with a density $f_m(t_1,\ldots,t_m,x_1,\ldots,x_m) \equiv f_m(x_1,\ldots,x_m)$, then the optimal forecast for r steps is the conditional average:

$$\begin{pmatrix} \hat{\xi}_{n+1} \\ \vdots \\ \hat{\xi}_{n+r} \end{pmatrix} = \mathrm{E}(\xi_{n+1}, \xi_{n+2}, \ldots, \xi_{n+r}\ \xi_1 = x_1, \ldots, \xi_n = x_n) = \boldsymbol{g}(x_1, x_2, \ldots, x_n), \qquad (1)$$

$$\xi_j = \xi(t_j).$$

The coordinates of the conditional mean are calculated by the formula:

$$\hat{\xi}_{n+j} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} y_j p(y_1, y_2, \dots, y_r\ x_1, \dots, x_n) dy_1 \dots dy_r\ ,$$

where the conditional density

$$p(y_1, y_2, \dots, y_r\ x_1, \dots, x_n) =$$

$$= f_{n+r}(x_1, x_2, \dots, x_n, y_{n+1}, \dots, y_{n+r}) \times (f_n(x_1, x_2, \dots, x_n))^{-1},$$

$$f_n(x_1, x_2, \dots, x_n) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_{n+r}(x_1, x_2, \dots, x_n, z_1, \dots, z_r) dz_1 \dots dz_r\ .$$

Optimality means that for random vectors

$$\xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}, \quad \eta = \begin{pmatrix} \xi_{n+1} \\ \vdots \\ \xi_{n+r} \end{pmatrix}$$

and the fair ratio

$$\mathrm{E}\eta - g(\xi)^2 = \min_{h} \mathrm{E}\eta - h(\xi)^2\ ,$$

where the function g is defined by the equality (1), $\|\cdot\|$ is a Euclidian distance in R^n .

As a rule, the density of the distribution f_m is unknown, and to calculate the forecast, it is necessary to create a stochastic time series model $\{x_1, x_2, \dots, x_n\}$, that is, to construct a description of the relationship between the values of the series based on previous observations.

This model can be described

$$\xi_k = \Psi(\xi_1, \dots, \xi_{k-1}, \varepsilon)\ ,$$

where $\{\xi_1, \xi_2, \dots, \xi_n\}$ are the sequence of random variables for which the observation x_k is assigned by some value ξ_k , and ε is a random vector. Note that standard models do not always take into account the specifics of observations and the corresponding forecast turns out to be far from reality (as shown in the example below).

The forecast constructions based on traditional time series models are discussed in detail in the reference [1] .

The preferred option of forecasting is a situation where the time series admits a Gaussian model, i.e. $\begin{pmatrix} \xi \\ \eta \end{pmatrix} \sim \aleph(0, Q)$, where the matrix Q is depicted in the form of

$$Q = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

and block elements A, B, C, D are determined by the ratios:

$$a_{jk} = E\xi_j \xi_k\ ,\ \ 1 \le j,\ k \le n \Leftrightarrow \xi \sim \aleph(0; A);$$

$$d_{jk} = E\eta_j \eta_k \equiv E\xi_{n+j}\xi_{n+k}\ ,\ \ 1 \le j,\ k \le r \Leftrightarrow \eta \sim \aleph(0; D)$$

and the elements of matrix $B, C,\ B = C^*$ are represent the mutual covariance of ξ and η, so

$$b_{jk} = E\xi_j\xi_{n+k} \ .1 \le j \le n, 1 \le k \le r;$$

$$c_{jk} = E\xi_{n+j}\xi_k \ ,\ 1 \le j \le r, 1 \le k \le n.$$

In this case, the optimal forecast is linear and is determined by the formula:

$$\hat{\eta} = \mathrm{E}(\eta \,|\, \xi) = CA^{-1}\xi\ . \qquad (2)$$

In particular, the sample $\{x_1, \ldots, x_n\}$ can be considered the values of the fractional Brownian motion $B_H(t)$, which is defined as a Gaussian random process with characteristics

$$\mathrm{E}B_H(t) = 0, \quad \mathrm{E}B_H(t)B_H(s) = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H})\ ,$$

where $0 < H < 1$ is called a Hurst exponent.

The properties of fractional Brownian motion are discussed in [2]. The experience with temporal data shows that identifying a series as fBm values is a rather rare phenomenon.

TIME SERIES NORMALIZATION METHOD

The idea of the method consists of transforming the original data $\{x_1, x_2, \ldots, x_n\}$. into Gaussian $\{u_1, u_2, \ldots, u_n\}$ ..

In some cases, the original data can be converted to the fractional Brownian motion values.

Let there be a continuous one-to-one mapping

$$\varphi: R^n \to R^n\ ,\ \varphi(x_1, x_2, \ldots, x_n) = (u_1, u_2, \ldots, u_n)$$

such that the increments of the transformed data

$$z_1 = u_1,\ z_2 = u_2 - u_1, \ldots, z_n = u_n - u_{n-1}$$

form a stationary sequence (i.e., they admit a stationary model).

The statistics

$$\sigma^2 = \frac{1}{n}\sum_{k=1}^{n} z_k^2 \quad \text{and} \quad \theta = \frac{1}{n-1}\sum_{j=1}^{n-1} z_j z_{j+1}$$

are consistent estimates of the variance of the increments and their one-step covariance ([3, 4]).

Then consistent estimates of the correlation coefficient and Hurst parameter are calculated by the formulas:

$$\rho = \frac{\theta}{\sigma^2}\ , \quad H = \frac{1}{2} + \frac{ln(\rho + 1)}{2\,ln 2}\ .$$

The necessary conditions for the hypothesis of "data $\{u_1, u_2, \ldots, u_n\}$ form the values of fractional Brownian motion" is the fulfillment of the limiting relations for the statistics A_n, B_n, D_n, F_n, which are defined by the following relations:

$$A_n = \frac{1}{n}\frac{1}{\sigma^4}\sum_{k=1}^{n} u_{k-1} z_k^3 \to -\frac{3}{2}, \quad H \in \left(0; \frac{1}{2}\right);$$

$$B_n = \frac{1}{n^{1+H}}\frac{1}{\sigma^5}\sum_{k=1}^{n} u_{k-1}^2 z_k^3 \to 3\eta, \;\; \eta \sim \aleph\left(0; \frac{1}{2H+2}\right), \;\; H \in \left(0; \frac{1}{2}\right);$$

$$D_n = \frac{1}{n^{2H}}\frac{1}{\sigma^4}\sum_{k=1}^{n} u_{k-1} z_k^3 \to \frac{3}{2}\varsigma^2, \;\; \varsigma \sim \aleph(0;1), \quad H \in \left(\frac{1}{2}; 1\right);$$

$$F_n = \frac{1}{n^{H}}\frac{1}{\sigma^3}\sum_{k=1}^{n} z_k^3 \to 3\varsigma, \quad H \in \left(\frac{1}{2}; 1\right).$$

The proof of the limit relations is contained in [5].

The standard algorithm for testing the specified hypothesis using known H and σ is a following: let us assume that the hypothesis is fulfilled and we set the significance level α with comparing the value of the statistic with the tabular value β where $F(\beta) = 1 - \alpha$.

In particular, for the marginal distribution function of statistics D_n $(H > 0{,}5)$:

$$F(x) = 2\Phi\left(\sqrt{x\frac{2}{3}}\right) - 1, \text{ where } \Phi \text{ is a Laplace function.}$$

Corresponding to the level of significance $\alpha = 0{,}05$ the critical value $\beta = 6$, and the hypothesis is accepted if $0 < D_n < 6$.

If the hypothesis $u_k = B_H(t_k)$ is true, then the optimal (linear) forecast for $\{u_{n+1}, u_{n+2}, \ldots, u_{2n}\}$ in *n* steps is calculated by the formula (2):

$$\begin{pmatrix} \hat{u}_{n+1} \\ \vdots \\ \hat{u}_{2n} \end{pmatrix} = CA^{-1}\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}, \quad Q = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

$$q_{jk} == (j^{2H} + k^{2H} - |j-k|^{2H}), \;\; 1 \le j, \; k \le 2n,$$

and the forecast of the primary variables

$$(\hat{x}_{n+1}, \ldots, \hat{x}_{2n}) = \varphi^{-1}(\hat{u}_{n+1}, \ldots, \hat{u}_{2n}).$$

Note that the choice of transformation φ is a rather cumbersome procedure.

Let us consider another normalization method — the one-dimensional transformation of the real data $u_k = \varphi(x_k)$.

$$\varphi(x) = \Phi^{-1}(F_\xi(x)), \qquad \textbf{(3)}$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \exp\left\{-\frac{z^2}{2}\right\} dz$, F_ξ is a distribution function corresponding to the sample $\{x_1, x_2, \ldots, x_n\}$.

Thereby, $\{u_1,\ldots,u_n\}$ is a sample from the general population $\aleph(0,1)$ and the proposed procedure requires stationarity in the narrow sense of the real data and requires an assessment of their distribution. The sample size should not be large to prevent the detection of non-stationarity.

Actually, $\{x_1,x_2,\ldots,x_n\}$ represents a sample of different random variables $\{\xi(t_1),\xi(t_2),\ldots,\xi(t_n)\}$ with the same distribution F. This disadvantage is partially compensated by the dependence of the obtained data, which is determined by their correlation matrix.

Let us note the following property of the transformation.

Proposition. Let φ is a strictly increasing function $\xi(t)$ is a stationary process in the narrow sense, i.e.

$$P\{\xi(t_1)<x_1,\ldots,\xi(t_n)<x_n\}=F(x_1,\ldots,x_n,t_2-t_1,\ldots,t_n-t_{n-1}).$$

Then the process $\eta(t)=\varphi(\xi(t))$ is also stationary in the narrow sense.

Proof. Under the condition of stationarity

$$P\{\xi(t_1)<x_1,\ldots,\xi(t_n)<x_n\}=F(x_1,\ldots,x_n,t_2-t_1,\ldots,t_n-t_{n-1}).$$

Distribution of the process $\eta(t)$:

$$P\{\varphi(\xi(t_1))<x_1,\ldots,\varphi(\xi(t_n))<x_n\}=P\{\xi(t_1)<y_1,\ldots,\xi(t_n)<y_n\}=$$
$$=F(y_1,\ldots,y_n,t_2-t_1,\ldots,t_n-t_{n-1}),\quad y_j=\varphi^{-1}(x_j)$$

also satisfies the definition of stationarity.

Let us formulate a modeling and forecasting algorithm using transformation (3).

1. Checking the data for stationarity, for example, using the Dickey–Fuller criterion and determining the size of the training sample (in the case of checking the adequacy of the model, determine the size of the training plus the predicted sample).

2. Estimation of the distribution function F_ξ of random value ξ by the sample $\{x_1,\ldots,x_n\}$.

3. Construction of the sample $\{u_1,\ldots,u_n\}$. from a normal population $\aleph(0,1)$ by the formula

$$u_k=\Phi^{-1}(F_\xi(x_k)).$$

4. Calculation of the sample correlation coefficients

$$\rho_j=\frac{1}{n-j}\sum_{k=1}^{n-j}u_k u_{k+j},$$

with consctruction of correlation matrix Q and defining the forecasting horizon r.

5. Construction of the forecast $\{\hat{u}_{n+1,\ldots,}\hat{u}_{n+r}\}$ of transformed data by the formula (2).

6. Calculating predicted values $\{\hat{x}_{n+1,\ldots,}\hat{x}_{n+r}\}$ of primary data according to the formula

$$\hat{x}_k=\varphi^{-1}(\hat{u}_k)\equiv F_\xi^{-1}(\Phi(\hat{u}_k)).\quad \hat{x}_{n+1},\ldots,\hat{x}_{n+r}.$$

EXAMPLE OF FORECAST CALCULATION

The following example illustrates the application of the proposed model:

The meteorological data files Precipitation–Florida Climate Center.

The sample size is determined by stationarity, which is tested using the Dickey–Fuller criterion.

In the given example, stationarity occurs in the interval $x_1 - x_{40}$; let us put the data into the training sample $x_1 - x_{30}$. The data values and their graph are shown in Table 1 and Fig. 1, respectively.

Table 1. The value of the series $\boldsymbol{x}_1 - \boldsymbol{x}_{40}$

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
3.03	**3.43**	**3.54**	**2.98**	**2.13**	**– 2.62**	**–0.61**	**–0.15**	**–2.15**	**–2.65**
x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}	x_{18}	x_{19}	x_{20}
–3.58	–2.13	–2.32	–2.43	–2.8	–2.42	–3.15	–2.62	–2.81	**–2.46**
x_{21}	x_{22}	x_{23}	x_{24}	x_{25}	x_{26}	x_{27}	x_{28}	x_{29}	x_{30}
–1.2	**–0.33**	**–1.81**	**–2.18**	**–0.2**	**0.6**	**3.07**	**5.48**	**6.34**	**8.81**
x_{31}	x_{32}	x_{33}	x_{34}	x_{35}	x_{36}	x_{37}	x_{38}	x_{39}	x_{40}
6.2	**4.04**	**2.86**	**1.53**	**0.702**	**1.7**	**3.72**	**5.33**	**6.27**	**3.32**

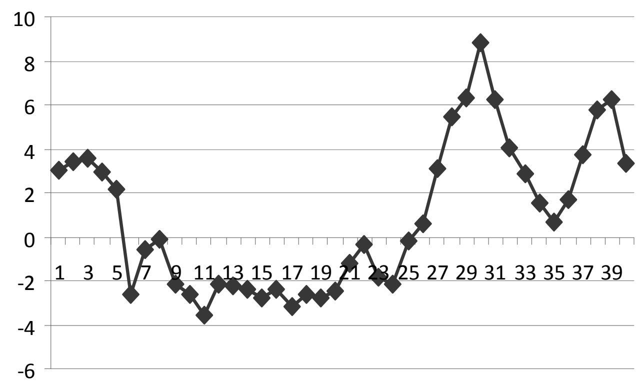


Fig. 1. Meteorological data $x_k = x(t_k)$

If we consider the given data as values of one random variable, then from the analysis of the sample $\{x_1, x_2, \ldots, x_n\}$ the hypothesis follows that this random variable ξ is distributed by Gumbel's law:

$$F_\xi(x) = \exp\left\{-\exp\left(\frac{\lambda - x}{\beta}\right)\right\},$$

with the moments:

$$E\xi = \lambda + \beta\gamma, \quad \gamma \approx 0{,}577 \text{ is a Euler's constant,}$$

$$\mathrm{E}\xi^2 = \int_0^\infty (\lambda - \beta \ln z)^2 e^{-z} dz = \frac{\pi^2}{6}\beta^2 + (\lambda + \beta\gamma)^2,$$

$$E\xi^3 = \int_0^\infty (\lambda - \beta \ln z)^3 e^{-z} dz = (\lambda + \beta\gamma)^3 + \frac{1}{2}\beta^2\pi^2(\lambda + \beta\gamma) + 2\beta^3\zeta(3),$$

$\zeta(s)$ is a Reamann zeta function, $\zeta(3) \approx 1{,}202$.

The values of sample moments:

$$\bar{x} = -0.037 \approx 0\,, \quad \overline{x^2} = \frac{1}{30}\sum_{k=1}^{30} x_k^2 = 10.19\,;$$

$$\overline{x^3} = \frac{1}{30}\sum_{k=1}^{30} x_k^3 = 33{,}76\,,$$

lead to estimates of the distribution parameters $\beta = 2{,}49$; $\lambda = -1{,}44$, so $F_\xi(x) = \exp\left\{-\exp\left(-\frac{x+1{,}44}{2{,}49}\right)\right\}$.

Density distribution graph

$$f_\xi(x) = \frac{1}{2{,}49}\exp\left(-\frac{x+1{,}44}{2{,}49}\right)F_\xi(x) \text{ shown in Fig. 2.}$$

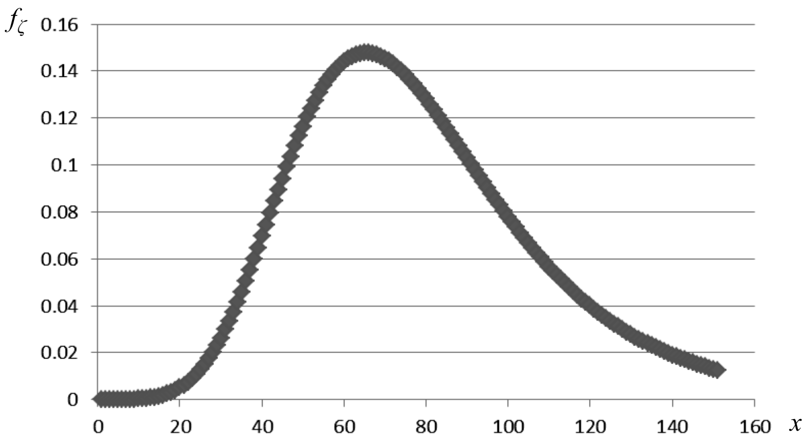


Fig. 2. Density distribution $f_\xi(x)$

Let us construct the transformed data using the formula

$$u_k = \Phi^{-1}\left(\exp\left\{-\exp\left(-\frac{x_k+1{,}44}{2{,}49}\right)\right\}\right) = \Phi^{-1}(y_k)\,,$$

where the values $\Phi^{-1}(y_k)$ are calculated using the Laplace function table. The result of the calculations is given in Table 2.

Table 2. Transformed data $u_1 - u_{40}$

u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8	u_9	u_{10}
1.02	**1.12**	**1.14**	**1.01**	**0.8**	**–0.84**	**–0.03**	**0.13**	**–0.63**	**–0.85**
u_{11}	u_{12}	u_{13}	u_{14}	u_{15}	u_{16}	u_{17}	u_{18}	u_{19}	u_{20}
–1.32	**–0.62**	**–0.65**	**–0.75**	**–0.92**	**–0.75**	**–1.09**	**–0.84**	**–0.92**	**–0.76**
u_{21}	u_{22}	u_{23}	u_{24}	u_{25}	u_{26}	u_{27}	u_{28}	u_{29}	u_{30}
–0.25	**0.07**	**–0.49**	**–0.64**	**0.11**	**0.37**	**1.03**	**1.56**	**1.72**	**2.13**
u_{31}	u_{32}	u_{33}	u_{34}	u_{35}	u_{36}	u_{37}	u_{38}	u_{39}	u_{40}
1.685	**–1.25**	**0.98**	**0.64**	**0.4**	**0.69**	**1.18**	**1.52**	**–1.71**	**1.09**

The evaluation of correlation coefficients

$$\rho_j = \frac{1}{30-j} \sum_{k=1}^{30-j} u_k u_{k+j} .$$

Leads to the results and for $k > 7$ the coefficients $\rho_k \approx 0$.

ρ_1	ρ_2	ρ_3	ρ_4	ρ_5	ρ_6	ρ_7
0.77	0.57	0.43	0.27	0.144	0.036	0.02

The elements of the matrix Q are determined by the ratios:

$$q_{jk} = \begin{cases} \rho_{|j-k|}, & |j-k| \le 7, \\ 0, & |j-k| > 7, \end{cases} \quad 1 \le j,\ k \le 37\ , \ \ \rho_0 = 1\ , \ \ a_{jk} = q_{jk}\ , \ \ 1 \le j,\ k \le 30 .$$

The forecast of transformed data in 7 steps is calculated by the formula (2)

$$\begin{pmatrix} \hat{u}_{31} \\ \vdots \\ \hat{u}_{37} \end{pmatrix} = CA^{-1} \begin{pmatrix} u_1 \\ \vdots \\ u_{30} \end{pmatrix} .$$

The forecasting results are shown in Table 3.

T a b l e 3 . The values of forecast in 7 steps

u_{31}	u_{32}	u_{33}	u_{34}	u_{35}	u_{36}	u_{37}
1.685	**1.25**	**0.98**	**0.64**	**0.4**	**0.69**	**1.18**
Forecast						
$\hat{u}_{31}$	$\hat{u}_{32}$	$\hat{u}_{33}$	$\hat{u}_{34}$	$\hat{u}_{35}$	$\hat{u}_{36}$	$\hat{u}_{37}$
1.682	**1.155**	**0.838**	**0.443**	**0.205**	**0.6**	**0.32**
$\hat{x}_{31}$	$\hat{x}_{32}$	$\hat{x}_{33}$	$\hat{x}_{34}$	$\hat{x}_{35}$	$\hat{x}_{36}$	$\hat{x}_{37}$
6.169	**3.595**	**2.267**	**0.848**	**0.08**	**1.396**	**0.44**

where the forecast of the initial data

$\hat{x}_{31} - \hat{x}_{37}$ is calculated by the formula

$$\hat{x}_k = \varphi^{-1}(\hat{u}_k) \equiv F_\xi^{-1}(\Phi(\hat{u}_k))\ , \quad F_\xi^{-1}(y) = -2{,}49 \ln(-\ln y) - 1{,}44 .$$

Let us compare the quality of forecasting using a model, which have constructed using the normalization method and four classical discrete time series models.

T a b l e 4 . Comparison of forecast quality

Time series values $x_{31} - x_{37}$							
Actual data	x_{31}	x_{32}	x_{33}	x_{34}	x_{35}	x_{36}	x_{37}
	6.2	**4.04**	**2.86**	**1.53**	**0.702**	**1.7**	**3.72**
Forecast data	**The results of forecast**						
	$\hat{x}_{31}$	$\hat{x}_{32}$	$\hat{x}_{33}$	$\hat{x}_{34}$	$\hat{x}_{35}$	$\hat{x}_{36}$	$\hat{x}_{37}$
Normalization method	**6.169**	**3.595**	**2.267**	**0.848**	**0.08**	**1.396**	**0.44**
ARMA	7.16	9.53	9.902	7.286	9.672	10.06	9.9
Point Forecast	6.703	5.163	3.89	2,833	1.927	1.142	0.434
GARCH	8.379	8.908	7.437	7.565	7.9	8.02	9.21
ARIMA	8.652	8.063	7.44	6.85	6.328	5.808	4.35

CONCLUSIONS

Given the stationarity of the time series, modeling using the normalization method, which is defined by the relation (3), provides higher forecast quality compared to traditional forecasting methods.

REFERENCES

1. P.I. Bidyuk, O.S. Menyaylenko, O.V. Polovtsev, *Forecasting methods*. Luhansk: Alma Mater, 2008, 604 p
2. Y. Mishura, "Stochastic Calculus for Fractional Brownian Motion and Related Processes," *Lecture Notes in Mathematics 1929*, Springer 2008, 393 p. doi: 10.1007/978-3-540-75873-0
3. K. Kubilius, Y. Mishura, K. Ralchenko, *Parameter Estimation in Fractional Diffusion Models*, vol. 8, 2018, 390 p. doi: https://doi.org/10.1007/978-3-319-71030
4. V.G. Bondarenko, "On some statistics of fractional brownian motion," *System Research & Information Technologies*, no. 1, pp. 131–138, 2021. doi: https://doi.org/10.20535/SRIT.2308-8893.2021.1.11
5. I. Nourdin, *Selected Aspects of fractional Brownian motion*. Milano: Springer, 2012, 124 p. doi: https://doi.org/10.1007/978-88-470-2823-4

Received 12.04.2024

INFORMATION ON THE ARTICLE

Viktor G. Bondarenko, ORCID: 0000-0003-1663-4799, Educational and Research Institute for Applied System Analysis of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine, e mail: bondarenvg@gmail.com

Valeriia V. Bondarenko, The University of the Littoral Opal Coast, France, e-mail: valeria_bondarenko@yahoo.com

ПРОГНОЗУВАННЯ ЧАСОВОГО РЯДУ ЗА МОДЕЛЛЮ НОРМАЛІЗАЦІЇ / В.Г. Бондаренко, В.В. Бондаренко

Анотація. Запропоновано емпіричні конструкції моделей часового ряду за схемою зведення первинних даних до нормально розподілених. Метою такого методу нормалізації є побудова оптимального прогнозу, який для оновлених даних є лінійним, а прогнозовані первинні дані відновлюються через обернене перетворення. Розглянуто варіанти таких перетворень — зведення первинних даних до гаусівського фрактального броунівського руху та одновимірне перетворення з використанням строго монотонної функції. Обчислювальний експеримент на базі реальних даних, що допускають стаціонарну модель, підтверджує вищу якість прогнозу методом нормалізації порівняно з традиційними моделями.

Ключові слова: оптимальний прогноз, стохастична модель, оцінювання параметрів, фрактальний броунівський рух.

UDC 519.711
DOI: 10.20535/SRIT.2308-8893.2025.2.08

STRATEGY FOR ENSURING ASYMPTOTIC CONVERGENCE OF THE PROCESS OF NON-LINEAR ESTIMATION OF DYNAMIC OBJECT PARAMETERS

**A. SILVESTROV, M. OSTROVERHOV, L. SPINUL,
O. KHALIMOVSKYY, H. VESHCHYKOV**

Abstract. The article considers a step-by-step strategy of sequential use and adjustment of a parallel model to an object of identical structure with orthogonal operators, a series-parallel model to an object with the connection of operators of a certain type for orthogonal approximation in order to obtain asymptotically unbiased estimates of coefficients of a structurally identical to a dynamic object of a mathematical model under conditions of noise of measurements of the initial variable of the identification object and non-convexity of the proximity functional of the initial variables of the object and the model in a space of coefficients of the object's mathematical model. Structural diagrams of each stage of identification are given using refined parameters and the structure of the model of object. This algorithm was implemented to identify the parameters of the mathematical model of aircraft, provided that the sample of experiment data is limited and there is of the initial a significant range of deviations of state variables from the basic mode.

Keywords: non-linear estimation, identification, convergence of estimation algorithms, optimization.

INTRODUCTION

An important place in the problems of non-linear programming (identification [1–8]) is occupied by the form of a proximity functional of initial variables of an object and a model, whose physical parameters are optimized at the extremum (usually the minimum) of the variable functional. This functional is simultaneously a function of physical parameters of the model being optimized. If the "input (*x*) – output (*y*)" mapping of the object can be represented by the following differential equation:

$$a_0 \frac{d^n y}{dt^n} + a_1 \frac{d^{n-1} y}{dt^{n-1}} + \ldots + a_{n-1} \frac{dy}{dt} + a_n y(t) =$$

$$= b_0 \frac{d^m x}{dt^m} + b_1 \frac{d^{m-1} x}{dt^{m-1}} + \ldots + b_{m-1} \frac{dx}{dt} + b_m x(t)\,, \qquad (1)$$

where a_i, $i = \overline{0,n}$; b_j, $j = \overline{0,m}$ are parameters to be determined using entries $y(t_k)$, $x(t_k)$, $k = \overline{1,N}$, considering that $y(t_k)$ is measured with uncorrelated noises $\zeta(t_k)$; then desired coefficients a_i of the model (1) enter the equation of error $\varepsilon(t_k)$ between $y(t_k)$ and output variable $y_M(t_k)$ non-linearly, since they are in the denominator of the model operator $W_M(p)$:

$$\varepsilon(t) = y(t) - W_M(p)x(t),$$

where $W_M(p) = \dfrac{b(p)}{a(p)}$, $p = \dfrac{d}{dt}$, $b(p) = b_0 p^m + b_1 p^{m-1} + ... + b_m$, $a(p) = a_0 p^n + + a_1 p^{n-1} + ... + a_n$.

Accordingly, parameters a_i non-linearly enter mean square $\overline{\varepsilon^2}$ of error $\varepsilon(t_k)$ (both for the functional of $y(t_k)$, $y_M(t_k)$ and the functions of parameters a_i, b_j), violating the elliptic form of dependence $\overline{\varepsilon^2}$ on deviations Δa_i of estimates a_i from their optimal values a_i^*:

$$(a_i^*, b_j^*) = \arg\min_{a_i, b_j} \overline{\varepsilon^2}. \tag{2}$$

Violation of ellipticity and strict convexity of function $\overline{\varepsilon^2}(a_i, b_j)$ (the "ravine" effect) significantly complicates the recurrent process of convergence of parameters a_i, b_j with their optimal values a_i^*, b_j^* that satisfy the condition (2).

Furthermore, the form of function $\overline{\varepsilon^2}(a_i, b_j)$ significantly depends on the bandwidth of input signal $x(t)$, because $\overline{\varepsilon^2}$ is simultaneously a functional of $y(t)$, $y_M(t)$, which, in turn, depend on input stimulus $x(t)$. For the given $x(t)$ and the structure of the model (1), the key to successful optimization of the relaxation process for convergence of the model (1) parameter estimates with their optimal values (2) can be a strategy of using a set of different models and a series of their connection and identification, which ensures convexity and ellipticity $\overline{\varepsilon^2}(a_i, b_j)$.

FORMULATION OF THE PROBLEM

While using an object-parallel:

- model with a structure identical to the object (1);
- model with a series-parallel connection of operators $\dfrac{a(p)}{c(p)}$ and $\dfrac{b(p)}{c(p)}$ to

the object (where $c(p)$ is a filter of the degree *p* exceeding the degree *n* of polynomial $a(p)$, which ensures the correctness of differentiation operation $y(t)$;

- as well as model with orthogonal parameters connected in parallel to the object; it is necessary to organize the sequence of their connection and adjustment in such a way as to ensure strict convexity and ellipticity of proximity functions of the object and the model being optimized, and thus, the convergence of the relaxation process for identification of the model (1) coefficients under the condition (2).

Strategy for implementing the condition of guaranteed convergence of model (1) parameters with their optimal values (2)

The strategy consists of four steps, where Steps 2, 3, 4 are to be repeated until condition (2) is met.

Step 1. Identification (Fig. 1) of weight coefficients β_i of operators $W_i(p)$ of the model (3).

An equation of the model parallel to the object:

$$y_M(t) = \sum_{i=1}^{n} \beta_i \{W_i(p) \cdot x(t)\} = \sum_{i=1}^{n} \beta_i \varphi_i(x(t)), \quad (3)$$

where $\varphi_i(x) = W_i(p)x(t)$ are linearly independent functions of time *t*.

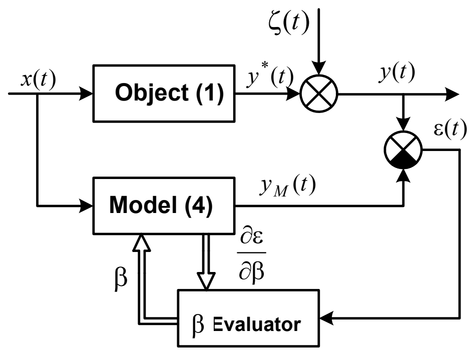


Fig. 1. Approximation of the "input-output" mapping (1) of the object in the model (3)

Operators $W_i(p)$ transform input stimulus $x(t)$ into a system of linearly independent functions $\varphi_i(x(t))$. If $x(t)$ is close to "white noise", then, using, for instance, Lager functions as these operators

$$W_i(p) = \frac{(p-\gamma)^i}{(p+\gamma)^i}, \quad (4)$$

(where γ is the operator parameter), we obtain a system of mutually orthogonal functions $\varphi_i(x(t))$. Linear independence, and even more so, orthogonality $\varphi_i(x(t))$ guarantee strict convexity and ellipticity of mean square $\overline{\varepsilon^2}$ of error $\varepsilon(t_k)$ [1], [2]. In the "off-line" mode, the determination of the model (3) coefficients β_i can be one-step, if we have a data sample $x(t_k)$, $y(t_k)$, $k = \overline{1, M}$. Then estimates $\hat{\beta}_i$ of coefficients β_i are determined by the least squares method (LSM) under the condition

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \frac{1}{M} \sum_{k=1}^{M} \varepsilon^2(k),$$

where $\varepsilon(k) = y(k) - \sum_{i=1}^{n} \beta_i \varphi_i(x(t))$.

Meaning, $\hat{\beta} = (\varphi^T \varphi)^{-1} \varphi^T Y = (\varphi^T \varphi)^{-1} \varphi^T (Y^* + \zeta)$.

LSM estimation $\hat{\beta}$ of the vector of coefficients β_i, $i = \overline{1, n}$ will be unbiased, since noise $\zeta(t)$ is uncorrelated with $\varphi_i(t)$, and if $\zeta(t)$ is a Gaussian "white noise", then estimate $\hat{\beta}$ will have minimal variance. If $x(t)$ is a Gaussian "white noise”, and operators $W_i(p)$ are type (4), then matrix $\varphi^{\mathrm{T}}\varphi$ will be diagonal, and each component $\hat{\beta}_i$ of vector $\hat{\beta}$ is determined independently:

$$\hat{\beta}_i = \frac{\sum_{k=1}^{M} \varphi_i(k)\, y(k)}{\sum_{k=1}^{M} \varphi_i^2(k)} .$$

In the “on-line” mode, the recurrent process of adjusting coefficients β_i of the model (3) is done through the gradient procedure:

$$\frac{d\hat{\beta}_i}{dt} = \lambda_i \varepsilon(t) \varphi_i(t) .$$

If λ_i is limited, the process of approximating estimate $\hat{\beta}_i(t)$ to the optimal stationary value is exponential, which achieves an exponentially weighted averaging of the current value $\varepsilon^2(t)$.

Resulting from operation of the system (Fig. 1) in the first step, with the limited dimensionality n of the model (3), we have a close approximation of mapping $x(t)$ onto $y^*(t)$ in the model (3) (non-parametric identification). In this case, the structures of the desired mapping (1) and the model (3) are different, but y_M no longer contains noise $\zeta(t)$.

Step 2. Approximation of mapping $x(t)$ onto $\hat{y}_M(t)$ that uses the series-parallel model (Fig. 2) of the equation (1).

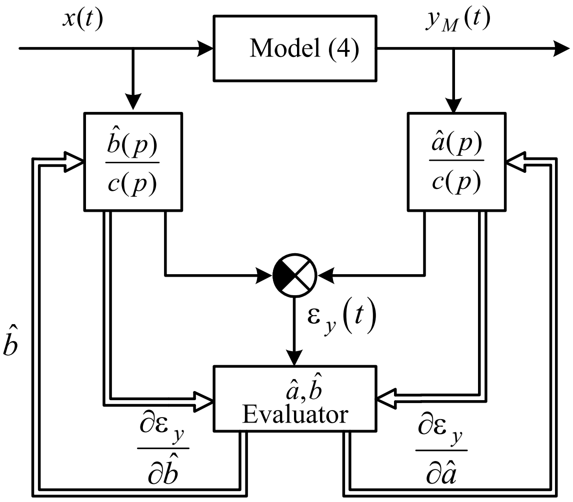


Fig. 2. Approximation of mapping $x(t)$ onto $\hat{y}^*_M(t)$ in the model (1)

Total error equation:

$$\varepsilon_y(t) = \frac{\hat{a}(p)}{c(p)} y_M(t) - \frac{\hat{b}(p)}{c(p)} x(t) , \tag{5}$$

where $c(p)$ is a filter polynomial with degree exceeding polynomial degree $a(p)$; polynomial structures $\hat{a}(p)$ and $\hat{b}(p)$ are identical to the polynomial structures $a(p)$ and $b(p)$ of the object model (1). Then, depending on the “online” or “off-line” modes of minimizing the mean square of total error $\varepsilon_y(t)$, by means of the adaptive circuit, the least squares method, or the gradient method, respectively, the optimal values of coefficients $\hat{a}_i$, $i=\overline{1,n}$, $\hat{b}_j$, $j=\overline{1,m}$ are calculated. Absence of “noise” in signal $\hat{y}_M(t)$, and linearity of dependence $\varepsilon_y(t)$ on coefficients $\hat{a}_i$, $\hat{b}_j$ being adjusted guarantee that their estimates are obtained. However, their values are not yet true values a_i, b_j of the object model (1). This is due to the proximity of mapping $y^*(t)$ in the model (3).

Step 3. Approximation of mapping $x(t)$ onto $y^*(t)$ in the model, in the form of a composition of the series connection of model $\frac{\hat{b}(p)}{\hat{a}(p)}$ obtained in the previous step, and the orthogonal model (3) (Fig. 3).

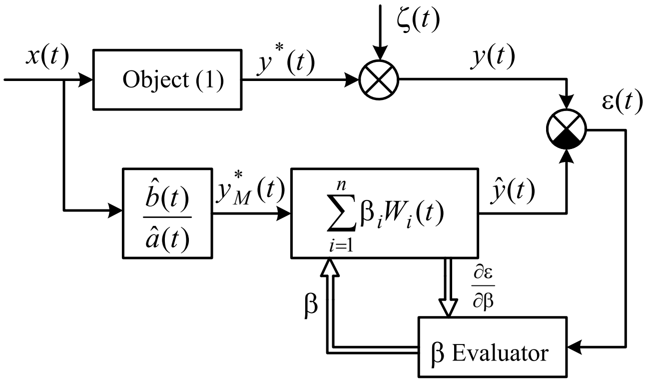


Fig. 3. Structure of the identification system in Step 3

In this step, an inaccuracy of mapping $x(t)$ onto $y^*(t)$ in the model $\frac{\hat{b}(p)}{\hat{a}(p)}$ is compensated by adjusting coefficients β_i of model $\sum_{i=1}^{n}\beta_i W_i(p)$, which is turned on in series with model $\frac{\hat{b}(p)}{\hat{a}(p)}$ (Fig. 3). The adjustment process β_i is similar to the process in the Step 1 (as per LSM in “off-line” mode, or gradient method in “on-line” mode). Now, however, model $\sum_{i=1}^{n}\beta_i W_i(p)$ should map not the mapping $x(t)$ onto $y^*(t)$, but only mapping $y_M^*(t)$ onto $y^*(t)$, which is much simpler, since $y_M^*(t)$ is already more or less close to $y^*(t)$.

Step 4 repeats Step 2 (Fig. 2), but for mapping $x(t)$ onto $y^*(t)$ clarified in the model $\sum_{i=1}^{n}\beta_i W_i(p)$. Theoretically, if in Step 4 n of the model $\sum_{i=1}^{n}\beta_i W_i(p)$ is unlimited, we will already get estimates $\hat{a}_i$, $\hat{b}_j$ of parameters a_i, b_j of the identification object’s accurate model (1), which will be unbiased by noise $\zeta(t)$.

In practice, if n is limited, Steps 2–4 are repeated, and $\hat{y}(t)$ gradually approaches $y^*(t)$ of the object, coefficients β_i of the model (3) approach zero, except for β_0, which approaches one with the unit operator $W_0(p)=1$, and coefficient estimates $\hat{a}_i$, $\hat{b}_j$ approach true values a_i, b_j of the model (1).

A flow chart of the strategy of using three types of models is presented in Fig. 4.

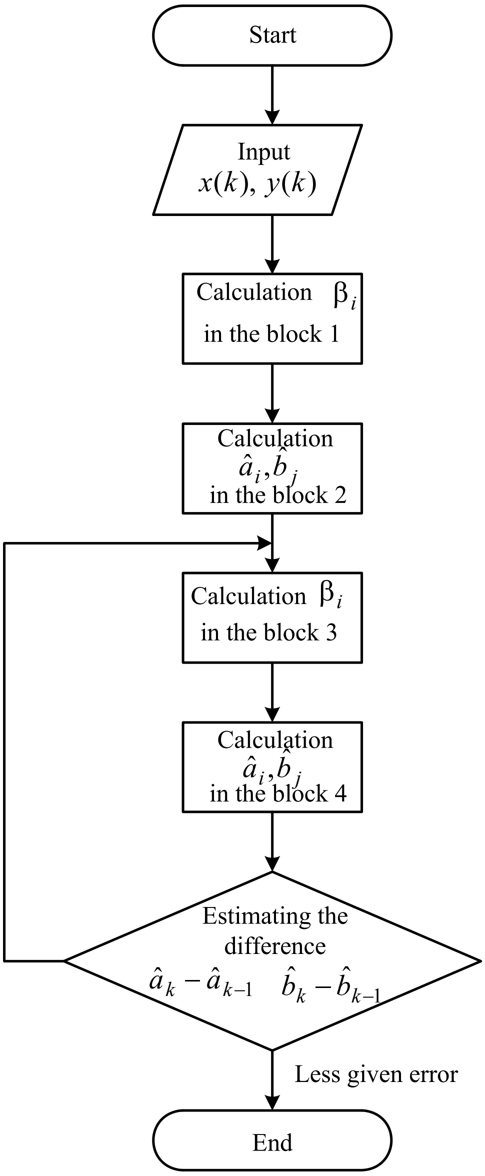


Fig. 4. Flow chart of the identification algorithm for model (1) parameters

We will consider the feasibility of using the proposed strategy on the example of identifying coefficients of the transfer function of an aircraft in longitudinal short-periodic movement [4]. The problem of identification lies in a concept of the transfer function being valid only under conditions of linearity and stationarity of the mapping of the control stimulus (deviation $\Delta\delta_h(t)$ of the altitude control) in the deviation of attack angle $\alpha(t)$ of the angle between the longitudinal axis of the aircraft and the direction of the oncoming air flow.

Short-periodic movement means movement in a short time interval during which factors not taken into account in model (1) hardly change. These include non-linearity, non-stationarity, speed changes, height, weight, and dimensions of the aircraft, etc. Therefore, the mathematical model of the aircraft movement will be more accurate, the smaller the time and deviation of the variables from a certain basic balancing mode are. If the time and range of deviations of the aircraft state variables from the basic mode tend to zero, the model (1) tends to being perfectly accurate. But in reality, the measurements of the aircraft state variables give not only accurate, but a random component of noise. Then, when the range of the variable deviations decreases, the "noise-to-signal" ratio increases, and, in a limited time interval, the variance of estimates of aircraft parameters will increase unacceptably. Desire to reduce impact of non-linearity on a linearized aircraft model leads to an increase in the "noise-to-signal" ratio and, as a result, an increase in the random component of the error in estimating the parameters of the mathematical model (MM) of the aircraft. This range increase leads to a shift in the linear MM parameters' estimates due to the influence of non-linearity of the aircraft characteristics. If it is impossible to fulfill other conditions (a large sample of experimental data and deviation range of the state variables from the basic mode), the condition of ellipticity and strict quadraticity of the error functional between $y(t)$ and $y_M(t)$, as a function of the optimized parameters of the aircraft MM, allow to successfully solve the problem of identifying parameters $\hat{a}_i$ and $\hat{b}_j$ of the MM on the level of their proximity to real physical values a_i and b_j of the accurate MM (1).

Let's suppose for small deviations from the basic longitudinal horizontal movement the aircraft MM looks as follows:

$$\frac{d^2y^*}{dt^2}+a_1\frac{dy^*}{dt}+a_0y^*(t)=bx(t)\,,$$

where $y^*(t)$ is the deviation of attack angle $\alpha(t)$, and $x(t)$ is the altitude control step deviation δ_h. Desired aerodynamic coefficients are $a_1=1\,c$, $a_0=3$, $b=0{,}5$.

Reaction $y^*(t)$ to a single step stimulus $x(t)$ is shown in Fig. 5.

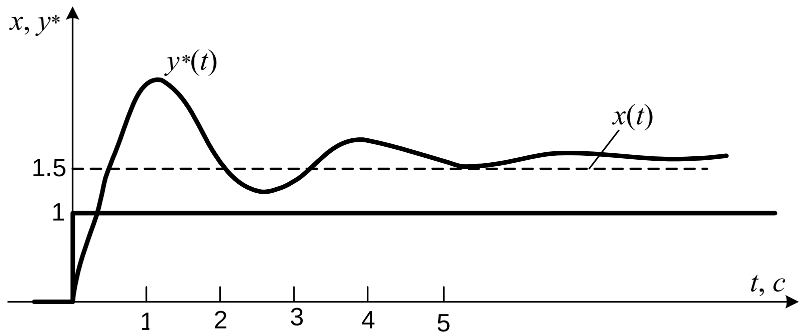


Fig. 5. Reaction $y^*(t)$ to step impact and $x(t)$

The test corresponds to real conditions of aircraft identification: observation time $T=5$ s, discretion step Δt in time was 0.02 s. Meaning, number M of discrete values t_k was 100. Measurements of attack angle $y(t)$ in 100 discrete-times t_k consist of the exact value $y^*(t_k)$ (Fig. 5), and adaptive noise in the form of Gaussian "white noise". Input stimulus $x(t)$ is measured accurately.

The computer modeling was carried out for various noise $y(t)$ to signal $y^*(t)$ ratios, therefore, due to a limited data sample and the presence of random noise, estimates $\hat{a}_0$, $\hat{a}_1$, $\hat{b}$ are random (Table). The estimates were obtained using the series-parallel model (Fig. 2, Step 2) and the three models strategy (Figs. 1–3, Steps 1–4). The model (3) uses three type (4) operators, while the model (5), as a filter $C(p)$, uses operator

$$W_f(p)=(p^3+5p^2+6p+7)^{-1},$$

which significantly smooths noise components $\zeta(t)$.

Table shows estimates $\hat{a}_0$, $\hat{a}_1$, $\hat{b}$ obtained through a one-step algorithm that uses only the series-parallel model (Step 2), and a multi-step (Steps 1–4) algorithm that uses the proposed strategy (Fig. 4).

The estimates, obtained by the one-step and multi-step algorithms

Noise/signal	$\hat{a}_0$		$\hat{a}_1$		$\hat{b}$	
	Step 2	Steps 1–4	Step 2	Steps 1–4	Step 2	Steps 1–4
0	3	3	1	1	0.5	0.5
0.5	3.15	3.1	1.01	1.1	0.517	0.516
1	2.69	3.15	0.71	1.14	0.6	0.52
2	1.56	2.82	0.16	0.96	0.49	0.48
3	1.39	3.22	0.02	0.95	0.53	0.58

Therefore, despite a random component of estimates $\hat{a}_0, \hat{a}_1, \hat{b}$, associated with a limited data sample and a significant noise $\zeta(t_k)$ to signal $y^*(t_k)$ ratio, in the dependences $\hat{a}_1$ and $\hat{a}_0$ on the noise-to-signal ratio, we can observe a pattern that has a statistically significant value, i.e. a significant decrease in estimates $\hat{a}_1$ and $\hat{a}_0$ with an increase in the noise level $\zeta(t_k)$ (Fig. 6).

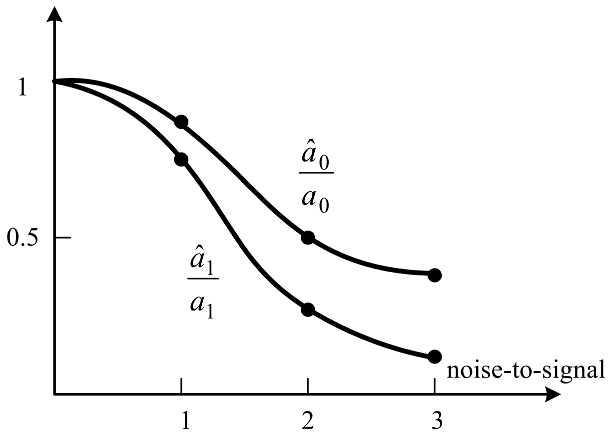


Fig. 6. Dependence $\frac{\hat{a}_1}{a_1}$ and $\frac{\hat{a}_0}{a_0}$ on the noise level $\zeta(t)$ in signal $y(t)$

To explain the effect of lower estimates $\hat{a}_i$ and $\hat{a}_0$, we will consider a structural diagram with a series-parallel model (Step 2), where an object is connected instead of a model (Fig. 7).

The LSM evaluator determines estimates $\hat{a}_0, \hat{a}_1, \hat{b}$ of parameters a_0, a_1, b under the condition of the minimum mean square $\bar{\varepsilon}_y^2$ of error $\varepsilon_y(t_k)$ [5; 6], i.e., under the following conditions:

$$\frac{1}{N}\sum_{k=1}^{N}\varepsilon_y(t_k)\frac{\partial\varepsilon_y(t_k)}{\partial\hat{a}_0} \equiv 0\,, \tag{6}$$

$$\frac{1}{N}\sum_{k=1}^{N}\varepsilon_y(t_k)\frac{\partial\varepsilon_y(t_k)}{\partial\hat{a}_1} \equiv 0\,, \tag{7}$$

$$\frac{1}{N}\sum_{k=1}^{N}\varepsilon_y(t_k)\frac{\partial\varepsilon_y(t_k)}{\partial\hat{b}} \equiv 0\,, \tag{8}$$

Total error:

$$\varepsilon_y(t_k) = \hat{a}_0\hat{y}(t_k) + \hat{a}_1\frac{d\hat{y}(t_k)}{dt} + \frac{d^2\hat{y}(t_k)}{dt^2} - \hat{b}\hat{x}(t_k)\,. \tag{9}$$

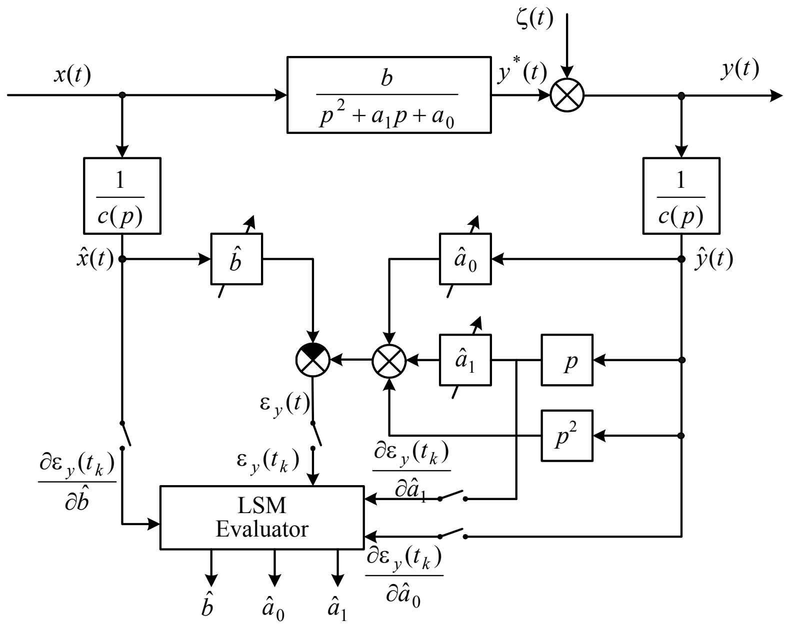


Fig. 7. Evaluation of object parameters using the series-parallel model

Error sensitivity functions for the relevant parameters:

$$\frac{\partial \varepsilon_y(t_k)}{\partial \hat{b}} = \hat{x}(t_k),$$

$$\frac{\partial \varepsilon_y(t_k)}{\partial \hat{a}_0} = \hat{y}(t_k) = \hat{y}^*(t_k) + \hat{\zeta}(t_k),$$

$$\frac{\partial \varepsilon_y(t_k)}{\partial \hat{a}_1} = \frac{d\hat{y}(t_k)}{dt} = \frac{d\hat{y}^*(t_k)}{dt} + \frac{d\hat{\zeta}(t_k)}{dt}.$$

In condition (7), considering expressions (8) and (9), there is no square of noise $\zeta(t)$. If (9) is weakly correlated with the inaccuracy of determining coefficients $\hat{a}_0$ and $\hat{a}_1$ in (9), estimate $\hat{b}$ of parameter *b* is almost unbiased (see Table, Step 2).

It is different for estimates $\hat{a}_0$ and $\hat{a}_1$. In equations (6), (7), there is a square of noise $\zeta(t)$ or its derivative. This leads to a bias in estimates $\hat{a}_0$ and $\hat{a}_1$ proportional to the noise level $\zeta(t)$. This is the main drawback of the series-parallel model (Fig. 7), which is eliminated by the proposed strategy (see Table, Steps 2–4).

CONCLUSION

In order to guarantee unbiased physical parameter estimates for mathematical models of dynamic objects in real conditions of limited data samples and dynamic ranges of state variables of an object, an effective strategy consists of a step-by-step use of a parallel model with orthogonal operators, a series-parallel model for approximating the orthogonal model, and a subsequent use of these models for a more accurate "input-output" mapping of the object and, consequently, an unbiased estimation of desired dynamic object parameters.

REFERENCES

1. A.A. Pavlov, V.V.Kalashnik, “Recommendations for choosing the zone of an active experiment for one-dimensional polynomial regression analysis,” *Visnyk NTUU “KPI”. Informatics, Operation and Computer Science*, no. 60, pp. 41–45, 2014.
2. K.Kh. Zelenskyi, V.M. Yhnatenko, and A.P. Kots, C*omputer methods of applied mathematics.* Kyiv: Ukrainskyi Litopys, 2000, 207 p.
3. A.M. Silvestrov, V.M. Sineglazov, *Theory of identification: textbook*. Kyiv: NAU, 2015, 477 p. Available: https://ela.kpi.ua/handle/123456789/30302
4. V.M. Sineglazov, M.K. Filiashkin, *Automated aircraft control systems: textbook*. Kyiv: NAU, 2003, 502 p. Available: https://er.nau.edu.ua/handle/NAU/ 52668\
5. V.V.Samsonov, A.M. Silvestrov, *Essays on the theory of identification*. Kyiv: NUFT, 2021, 166 p. Available: https://ela.kpi.ua/handle/123456789/30522
6. V.V. Samsonov, A.M. Silvestrov, and L.Yu. Spinul, *Multiple adaptive system of identification*. Kyiv: NUFT, 2018, 228 p. Available: https://ela.kpi.ua/ handle/123456789/30306
7. I. Levchuk, G. Manko, V. Tryshkin, and V. Korsun, *Theory and practice of control objects identification*: *monograph.* Dnipro: Ukrainian State University of Chemical Technology , 2019, 203 p.
8. Kamil Aida-zade, Yegana Ashrafova, “Determination of the parameters of dynamic objects, arbitrarily related by conditions,” *Journal of Industrial and Management Optimization*, vol. 20, issue 7, pp. 2477–2499, 2024. doi: 10.3934/jimo.2024010

Received 13.03.2024

INFORMATION ON THE ARTICLE

Anton M. Silvestrov, ORCID: 0000-0002-2511-5029, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Ukraine, e-mail: silvestrovanton@gmail.com

Mykola Ya. Ostroverkhov, ORCID: 0000-0002-7322-8052, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Ukraine, e-mail: n.ostroverkhov@hotmail.com

Liudmyla Yu. Spinul, ORCID: 0000-0002-4234-6072, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Ukraine, e-mail: spinul20@gmail.com

Oleksiy M. Khalimovskyy, ORCID: 0000-0003-3672-8530, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Ukraine, e-mail: o.khalimovskyy@ukr.net

Heorhii V. Veshchykov, ORCID: 0009-0002-0606-9765, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Ukraine, e-mail: veschikov98@gmail.com

СТРАТЕГІЯ ЗАБЕЗПЕЧЕННЯ АСИМПТОТИЧНОЇ ЗБІЖНОСТІ ПРОЦЕСУ НЕЛІНІЙНОГО ОЦІНЮВАННЯ ПАРАМЕТРІВ ДИНАМІЧНИХ ОБ’ЄКТІВ/

А.М. Сільвестров, М.Я. Островерхов, Л.Ю. Спінул, О.М. Халімовський, Г.В. Вещиков

Анотація. Розглянуто покрокову стратегію послідовного використання і налаштування паралельної до об’єкта моделі ідентичної структури з ортогональними операторами, послідовно-паралельної моделі до об’єкта з підключенням операторів певного типу для ортогональної апроксимації з метою отримання асимптотично незміщених оцінок коефіцієнтів структурно ідентичної до динамічного об’єкта математичної моделі в умовах зашумленості вимірів вихідної змінної об’єкта ідентифікації і невипуклості функціонала близькості вихідних змінних об’єкта і моделі в просторі коефіцієнтів математичної моделі об’єкта. Наведено структурні схеми кожного етапу ідентифікації з використанням уточнених параметрів і структури моделі об’єкта. Алгоритм реалізовано для ідентифікації параметрів математичної моделі літальних апаратів за умови обмеженості вибірки даних експерименту і значного діапазону відхилення змінних стану від базового режиму.

Ключові слова: нелінійне оцінювання, ідентифікація, збіжність алгоритмів оцінювання, оптимізація.

UDC 004.8: 004.421.2
DOI: 10.20535/SRIT.2308-8893.2025.2.09

THE CONCEPT OF INTELLIGENT TRAINING SYSTEM FOR UKRAINIAN SCHOOL FINAL STEM EXAM PREPARATION

A. KULIK, O. ZELENIAK, A. CHUKHRAY, O. PROKHOROV, O. YASHYNA, O. HAVRYLENKO, O. YEVDOKYMOV, A. TORZHKOV, O. ZAYARNYI

Abstract. The National Multi-subject Test has been prepared and conducted in Ukraine in online learning conditions for several years. Test results show a decline in schoolchildren's performance in mathematics. The article presents a prototype of an intelligent training system for solving mathematical problems that should become an accessible test preparation tool. The system provides a solution to a wide range of mathematical problems in a step-by-step mode. The system is developed in accordance with the principle of rational management by diagnosis, which implies the presence of many diagnostic models. It allows for deep diagnostics of student errors. Artificial intelligence tools will make it possible to implement individual recommendations for each student, taking into account their level of preparation and learning goals.

Keywords: intelligent tutor system, National multi-subject test, math problems.

INTRODUCTION

The traditional education system was formed in the conditions of gradually increasing volumes of information and focused on the formation of a certain critical amount of knowledge, on the basis of which the necessary skills were formed. In modern conditions of digitization and exponential increase in information load, even a highly qualified instructor is not able to effectively adapt the educational process to each pupil or student in the class (group), also taking into account the need for inclusive education.

Intelligent Tutoring Systems (ITS) are effective tools of solving the problem. At the same time, none of the developed ITS has yet approached the effect of individual learning described by the famous American psychologist B. Bloom: the average success rate of a student who learns individually can be better than the success rate of 98% of students who learn in a traditional way: one mentor for thirty students.

Thus, the development of ITS is relevant for improving the efficiency and quality of education, which allow to combine, within the requirements of the modern educational system, the existing methodical and pedagogical experience of leading mentors with an individual approach to the education of each of the pupils and students with different basic levels of training, motivation and peculiarities of perception information

The paper proposes the author's scientifically based viewpoint at the technology of creating ITS in exact sciences for preparation for the National multi-subject test (NMT).

LITERATURE REVIEW

The paper [1] proposes a taxonomy of artificial intelligence (AI) methods used in digital tools for teaching mathematics. The taxonomy consists of four categories that cover the entire range of such AI systems:

1. Information Extractors to denote AI technologies that take observations from the real world (test, audio, images) and transform them into a mathematical representation.
2. Reasoning Engine, which includes all software systems capable of automatically solving a mathematically formulated problem.
3. Explainers, i.e., the field of explanatory artificial intelligence research that deals with the development of artificial intelligence methods that produce interpretable models and interpretable solutions.
4. Data-driven Modeling, where methods of intelligent data analysis and machine learning are used to analyze this data and turn it into practical models.

In [2], an intelligent tutoring system for mathematics uses the author's ITSB tool with four modules — domain, learning, student, and user interface — coordinated by the training module. The system collects personal and academic data, tracks student progress, and adapts learning paths based on profiling and activity monitoring.

In [3], an intelligent system teaches fractions multiplication and division via adaptive dialogue, identifies errors in real time, and employs cognitive conflict, problem simplification, and representative learning.

In [4], an AI-based system improves mathematics education strategies in underprivileged regions by integrating rating scales, norms, improvement strategies, and an intelligent assessment program. It diagnoses learner performance in batches and automatically proposes targeted improvements.

Scientists of the National Aerospace University "Kharkiv Aviation Institute" proposed their own approach to the creation of intellectual educational programs. The article [5] discusses the technology of building intelligent computer programs for learning algorithms. The subject area of study is an algorithmic tasks which are characterized by properties of determinism, mass and efficiency. Developments are based on an approach to the rational control of objects in conditions of partial uncertainty. Proposed information technology based on an approach to the rational control of objects in conditions of partial uncertainty.

The process of step-by-step solution of algebraic equations is considered in the article [6]. On the basis of the signal-parametric approach to diagnostics of faults in dynamic systems the mathematical diagnostic models are created which allow detecting classes of errors by comparing the results of Student's calculations and the results of system calculations. The approach to the formalization of the generation of problem situations applicable to the development of tutoring programs consisting of many tasks is considered in [7]. The parametric generation method proposed in the article allows getting the large quantitative variations in task problem situations. Thereby, every learner will get a personal unique set of tasks.

In [8], OnlineMSchool offers math problems of varying complexity and online calculators that show step-by-step solutions.

In [9], IXL Learning provides interactive exercises, personalized recommendations, and real-time diagnostics for math and language skills.

In [10], Khan Academy delivers free, self-paced courses across various subjects, featuring thousands of videos, adaptive exercises, and progress tracking. Its paid AI tool, Khanmigo, uses ChatGPT technology to serve as a personal tutor for students and an assistant for teachers.

LearnBop [11] is a virtual system guiding children step-by-step through math problems, pinpointing gaps in foundational knowledge. Uses images, graphics, and video hints (via LearnZillion). Teachers can customize lessons, track progress, and assign interventions with the built-in LMS. Children enjoy the panda mascots during lessons.

Mangahigh [12] is a web platform offering math games, tutorials, and tests aligned with Common Core standards, covering fundamentals from counting to early algebra and geometry. Learning adapts to the learner's level, and wrong answers come with explanations. More practice variety could further enhance skill depth.

GeoGebra [13] is a free program enabling students to create, visualize, and manipulate mathematical models in algebra, geometry, and more. It has browser and mobile versions, with a vast library of ready-made resources. Though there's a learning curve, its interactive approach helps students explore and understand math concepts hands-on.

CueThink [14] introduces Thinklets: four-step solutions — Understand, Plan, Solve, and Revise. Students view teacher assignments, create videos showing their solutions, and share them with peers for feedback. Tools like markers and notebooks support each step, fostering collaboration and deeper problem-solving skills.

WeBWorK [15] is a free Perl-based system (launched in 1995) for online math homework, developed at the University of Rochester. Gives students immediate feedback, encouraging multiple attempts. Distinguishes itself by integrating LaTeX with Perl for flexible problem generation.

In turn, IMathAS [16] is an internet-based math assessment system generating algorithmic questions with numeric or symbolic answers. Includes learning-control tools, shows math and graphs accurately, and accepts simple calculator-style inputs..

WirisQuizzes is a tool for creating STEM assessments with equations, graphs, or text answers, automatically graded. Random parameters help prevent cheating. It integrates seamlessly with different systems, allowing dynamic question generation in real time.

Recently, many math-focused educational resources have begun using generative AI (e.g., ChatGPT). While machine learning has advanced significantly, including ChatGPT, it lacks a "true understanding" of math and can struggle with structured computation. Though ChatGPT may generate plausible explanations, it is often unreliable for precise answers. A solution involves connecting ChatGPT to Wolfram|Alpha, whose computational "superpowers" rely on Wolfram Language [17] for exact code and computations.

Another interesting study in the direction of the application of AI for mathematics was presented by the division of the Google company Deepmind. AlphaGeometry [18] solves complex geometry problems at near — Olympic gold medalist level, a notable breakthrough in AI. In testing with 30 Olympiad geometry problems, AlphaGeometry solved 25, compared to 10 by the previous best system and 25.9 by the average gold medalist. AI often struggles with complex geometry due to limited reasoning and training data. AlphaGeometry tackles this by merging a neural language model's predictive power with a rule-based deduction engine.

Thus, the possibilities of preparing for the NMT in mathematics are quite powerful, schoolchildren and teachers can choose the tools at their discretion and estimate the effectiveness of learning on trial tests that have been developed by the Ukrainian Center for Education Quality Evaluation. On the other hand, we will analyze the results of last year's testing of Ukrainian students within the NMT, using data from the official report [19].

The certification work contains tasks of various forms:

1. Tasks 1–15 with the choice of one correct answer (1 point).
2. Tasks 16–18 for matching, determining "logical pairs" (3 points).
3. Tasks 19–22 open form with a short answer (2 points).

The maximum number of points that could be obtained by correctly completing all the tasks of the certification paper in mathematics is 32.

The mathematics test was included in the main block of the NMT, the subjects of which were mandatory for all test participants. Fig. 1 shows the distribution of mathematics test participants by the number of test points scored. The diagram shows that the vast majority of participants scored less than 10 points, which is a very low indicator for the possibility of further study at a university.

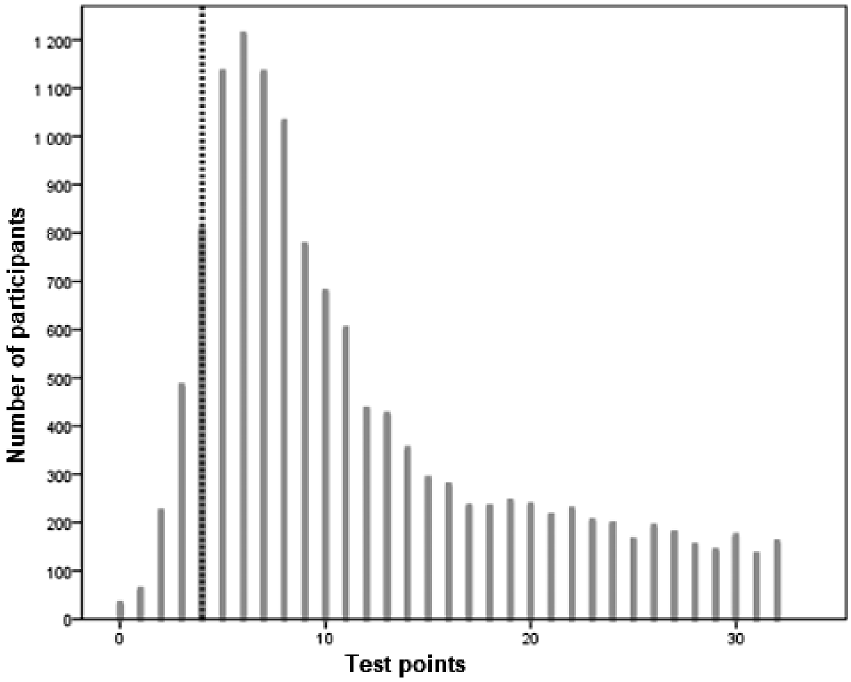


Fig. 1. Distribution of mathematics test participants by the number of test points scored

The share of competence tasks, namely: the task of constructing a pie chart according to a known distribution, calculating the percentage of a number, classifying a body formed as a result of the rotation of a geometric figure around an axis, and the combinatorial task of determining the number of options using combinations and multiplication rules — was 23 % of all test tasks.

The analysis of the results of the tasks of the certification work showed that even the tasks for checking the level of formation of basic skills and abilities and their application when solving standard problems cause difficulties for the participants. The number of such tasks was two-thirds of the total number of certification work tasks. For example, 48% of participants successfully coped with task 6 on the reproduction of facts (determining the zero of the function according to the given graph). Problem 20 from combinatorics caused the greatest difficulties, only 12.7% of those tested gave the correct answer.

The statistical indicators of the NMT tasks have slightly deteriorated compared to the corresponding indicators of the VET tasks of previous years. Thus, in 2020, more than two-thirds of the participants completed the task of determining the zero of the function according to the given figure, and in 2023, less than half. In 2020, 60% of the test subjects completed a short-answer task to check the financial literacy of participants, which involved finding percentages from a number, and in 2023, only 53.6% completed a similar task with choosing the correct answer from five offered [20].

The analysis of the statistical indicators of the certification work revealed a high distributive ability of the test tasks with an average score of 56.7, which made it possible to single out the participants prepared for study in higher education institutions.

The analysis of the results of the NMT in mathematics for the past year, as well as the tendency of deterioration of performance indicators in previous years, makes the task of improving the methodology of preparing for the NMT in mathematics with the help of intelligent computer training programs an urgent and important task.

THE LEARNING PROCESS FROM THE POINT OF VIEW OF CONTROL THEORY

There are three classical control principles: impact control, disturbance control, and deviation control. They allow you to achieve the goal in different ways.

Many destabilizing factors — well-being, mood, motivation, and a student's initial preparation — affect progress toward learning goals. We can't effectively manage these without identifying them through interaction with the student. A different approach, the principle of control by diagnosis, is needed to address these challenges.

Let's consider the application of various control principles in the educational process.

When teaching calculation skills (e.g., in math, physics, or chemistry), two types of teachers illustrate how control principles are applied: Schematically, such a learning process can be represented by the following diagram (Fig. 2).

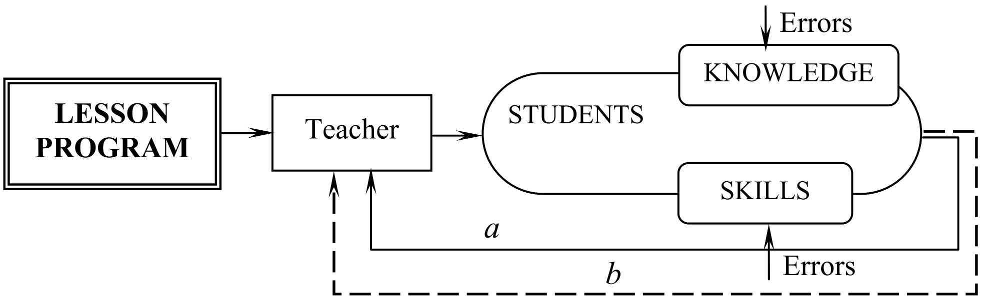


Fig. 2. Scheme of training control: *a* — regular; *b* — irregular

Teachers (Fig. 2, *b*) follow formal procedures and schedules without deeply checking each student's work, feedback is irregular, only at control points or at the end of class, quality and efficiency suffer because errors are not addressed quickly

Professional teachers (Fig. 2, *a*) provide ongoing, personalized feedback. They regularly identify errors, clarify misunderstandings, and strengthen student motivation. This "vector feedback," or continuous monitoring and correction, aligns with personalized training methods and achieves the highest efficiency and quality of learning.

High efficiency and significant quality of skill training are achieved with individual training. With such training, the teacher has the opportunity to identify "gaps" in knowledge, reasons for misunderstanding the theory, errors in the ability to apply theoretical knowledge in the practice of calculations, weak motivation and a number of other reasons for imperfect ability to solve calculation tasks, as well as to quickly increase the level of knowledge and the necessary skills to overcome the difficulties found in the student. The learning control process in this case is based on operational system diagnostics of the student's knowledge and skills (Fig. 3).

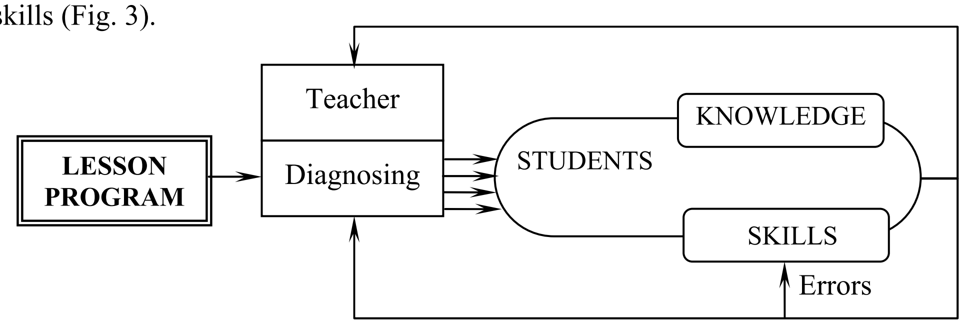


Fig. 3. Scheme of rational training control

In such a learning process, the teacher has more opportunities to diagnose the causes of errors made when solving the calculation task, as well as more opportunities to eliminate them with help the involvement of additional material that is more accessible to this student with the help of an insightful explanation of the causes of errors and ways to eliminate them. Diagnosing errors in this learning process is a highly intellectual type of teaching activity that allows you to identify weaknesses in the student's training in the ability to acquire new knowledge and use it in practical activities, in motivation.

Productive control of skill training is based on the use of the principle of diagnosis control [22].

Based on the results of the analysis of the above principles, and taking into account the work created by Professor A.S. Kulik approach to the rational control of objects in conditions of partial uncertainty was formed by the block diagram of the system of rational control of education with the help of ITS, depicted in Fig. 4.

As shown in figure the 6 main blocks in the learning control system with the help of ITS are the student diagnosis and modeling block (BDMO), as well as the learning action selection block (BBBV). Through the feedback channel, the data entered by the students in the IKOP are sent to the BDMO. In addition, as input actions on the BDMO, the necessary KZU and reference data are required in the ITS. BDMO affects BVOV by transferring the diagnosis, and even the model of the student.

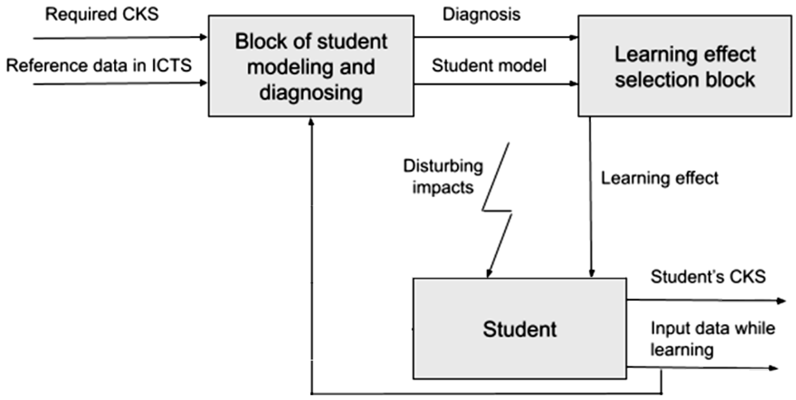


Fig. 4. Block diagram of the system of rational control of training with the help of ITS

PECULIARITIES OF THE COMPUTER IMPLEMENTATION OF THE PRINCIPLE OF CONTROL BY DIAGNOSIS

The principle of control by diagnosis has been applied in various projects focused on rational control of physical objects, using diagnostic models and knowledge bases on emergency behavior. Its computer-based implementation for training involves breaking down calculation tasks into stages, each producing intermediate numerical results for verification. Since mistakes can occur at any operation, each stage's result is compared with a reference outcome. Errors are then diagnosed by locating the incorrect operation and identifying the specific error. This step-by-step diagnostic process relies on a detailed solution model, making a production knowledge base and dichotomous tree structure ideal for implementing diagnostic control.

The consolidated structure of the curriculum, which implements the principle of diagnosis control, can be presented graphically in the following way (Fig. 5).

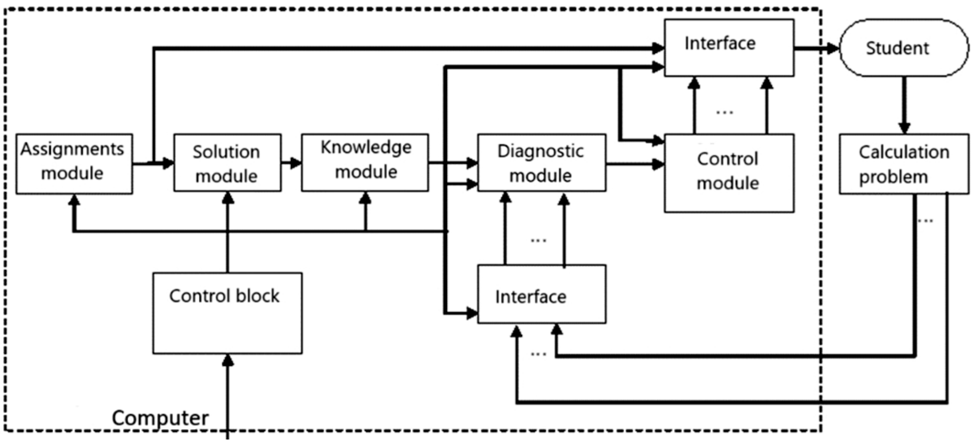


Fig. 5. Scheme of learning skills using the principle of diagnosis control

The scheme presents the functionally necessary software modules for the implementation of the principle of control based on the diagnosis. In the key module — the diagnostic module, a step-by-step program for diagnosing the results of student problem solving is implemented.

The module receives data from the knowledge base (machine solution) and the student's solution. After diagnosing the student's errors, it sends the results to the resource control module, which provides explanations and guidance for correction. The student then recalculates the problematic stage, iterating until their results match the correct solution. A control unit synchronizes all modules, and upon completion, generates performance indicators and an assessment of the student's ability. This iterative process arises due to insufficient preparation and the inevitable calculation errors.

A GENERAL APPROACH TO COMPUTERIZED LEARNING

The learning process is realized by a combination of internal and external cycles. The process of problem solving forms an internal learning cycle.

The student receives a task. He can immediately go to the answer or choose a step-by-step mode. In the step-by-step mode, he receives hints previously created by the teacher and implements his solutions. In this, it performs both calculations and symbolic or graphic transformations. If the answer is correct, it is highlighted in green and the student can proceed to the next task. If an incorrect solution is entered at the next step, the system issues the message "Think!" without specifying what the error is. The student can correct the error himself or ask for a hint. In the case of a repeated error, the input field is highlighted in red, visualizing the incorrect component of the solution. At the next error, built-in models are included to search for the cause of the error and generate an individual prompt. If the student could not complete the task, he can request a solution to this step. The correct solution is filled in by the system and highlighted in orange, after which the Student is prompted to proceed to the next step. The internal cycle is schematically presented in Fig. 6, 7.

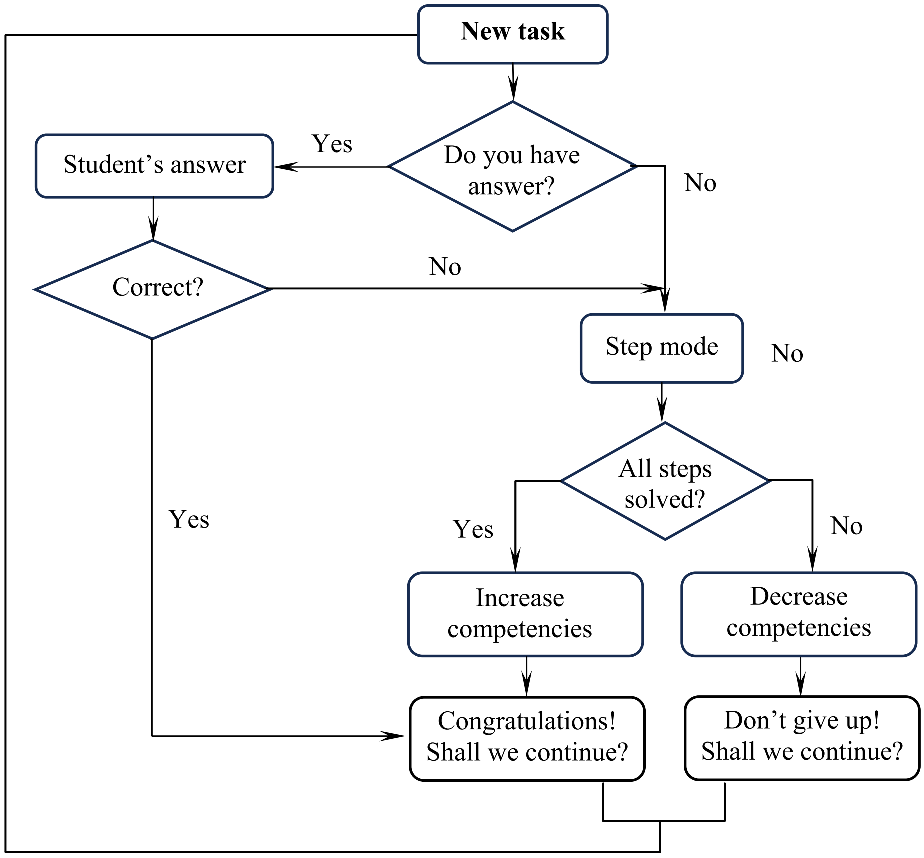


Fig. 6. Internal cycle of solving the task

The external cycle manages the learning process depending on the achieved competencies and the results of solving tasks. It can be presented schematically in the form of a Table 2.

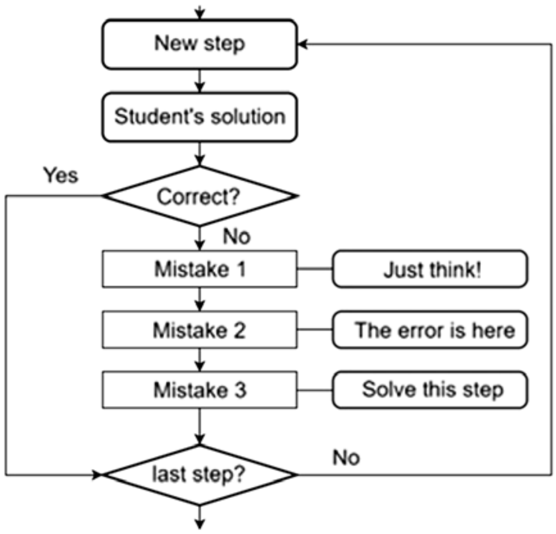


Fig. 7. Step mode

T a b l e 2 . External cycle of task selection

		Last task solved	
		Correct	Incorrect
Competencies level	High	Suggest a high difficulty problem or move on to the next section	Decrease the competency level and suggest the next task
	Low	Increase competency level and offer a more complex task	Decrease the competency level and: – offer work in demo mode; – recall theoretical material; – propose a simple task

The choice of the next task depends on the results of solving the previous one and previously achieved competencies. If the student successfully coped with the task, he is offered a more difficult task or a transition to the next chapter. If he made a mistake, and before that he had a high level, another task is offered. If he had a low level of competence and made mistakes again, he is offered a simpler task, as well as an opportunity to familiarize himself with the solution of tasks in a demonstration mode, that is, without evaluation. In addition, links to the theoretical material, which is desirable to repeat, are offered.

Based on the results of solving the problem, the student's level of competence is assessed on a probability scale (the probability of mastering this competence ranges from 0 to 1). Initially, it is equal to 0.5. Successful decisions increase this possibility, and mistakes decrease it. Achieving a level of 0.95 means mastering this competency.

Systematically modeling effective pedagogical activity, we will get three modes of operation of ITS: demonstration, training and test (Fig. 8). Thus, first the instructor explains the new material, demonstrating how new tasks for the students are solved, then calls the students to the blackboard and formulates tasks from the considered class of tasks, while helping the students in case of difficulties or wrong actions. Finally, to test the acquired knowledge and skills, the mentor gives homework or a control work, which the students solve independently, without prompts.

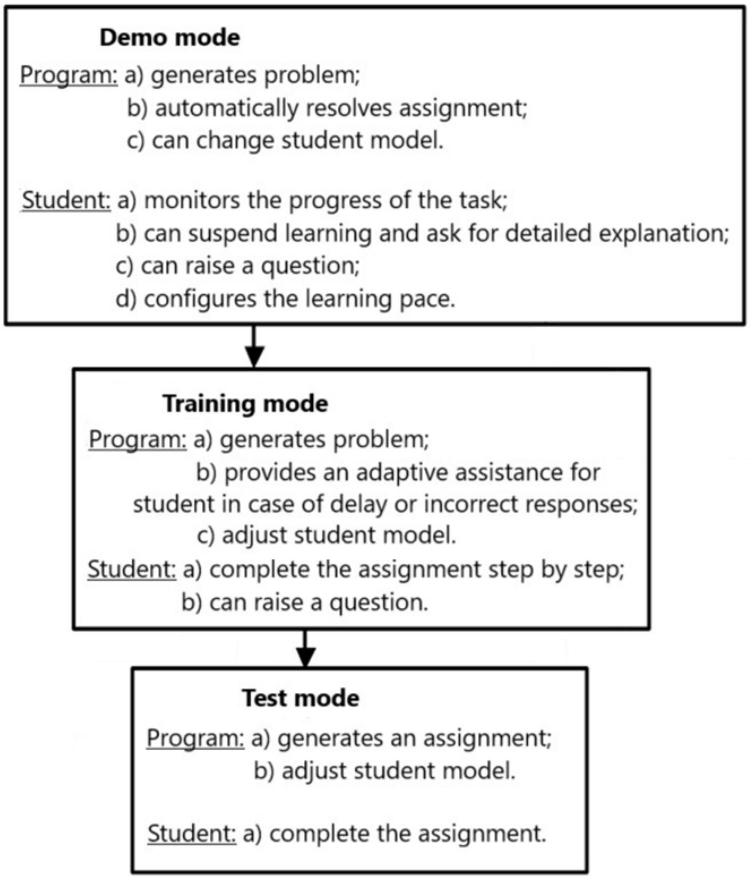


Fig. 8. Operating modes of education al programs

In the second mode, the task is formulated by the program, and the learner must complete it step by step. In case of mistakes or substitutions of the student, the ITS can give him a hint. The student can also ask questions to the program.

In the third mode, the task is formulated by the program, and the learner has to complete it. Unlike the previous mode, hints are not available. Also, the student does not have the opportunity to ask the program questions.

In order to develop stable skills for solving mathematical problems, the student must solve a number of problems of a certain class, which differ in conditions and numerical parameters. This increases the degree of assimilation of competencies, and also prevents memorization of answers to improve results when retaking tests. For this, it is necessary to ensure:

1) a set of meaningful statements for individual classes of mathematical problems (if possible);

2) the possibility of generating random numerical parameters for all classes of tasks.

To store problem templates in the system, the LATEX format is used — the data markup language and the TeX macro package, which is considered the de facto standard for preparing mathematical and technical texts for publication in scientific publications [109].

When the task template is entered, the teacher will have several fields assigned to different areas on the student's screen, which he can use to fill out in LATEX format. Numerical parameters are generated according to the scenario defined by the teacher.

For example, a hint template for one of the steps of a vector algebra task looks like this:

```
step_21_2.Text_Sample ="We got the equation \n\n"+ \
"$\overrightarrow{{AB}}({task21.xab}; {task21.yab};
{task21.zab}) = \overrightarrow{{(({task21.xb})- xa;
{task21.yb}) – ya; ({task21.zb}) - za)}}$ \n\n" + \
"Deduce from it equations for individual coordinate components in the form: \n\n"+ \
"$xab=xb-xa$, \n\n"+ \
"$yab=yb-ya$, \n\n"+ \
"$zab=zb-za$, \n\n"+ \
"where $\overrightarrow{{AB}}(xab; yab; zab)$ -
coordinates of the vector $\overrightarrow{{AB}}$"
```

After its interpretation by the program and generation of numerical parameters, we will get the prompt shown in Fig. 9.

We got the equation

$\overrightarrow{AB}(-3;8;1) = \overrightarrow{((7)-xa;(-2)-ya;(0)-za)}$

Deduce from it equations for individual coordinate components in the form:

$xab = xb - xa,$

$yab = yb - ya,$

$zab = zb - za,$

where $\overrightarrow{AB}(xab;yab;zab)$ - coordinates of the vector $\overrightarrow{AB}$

Fig. 9. A hint for a step-by-step solution of the problem with generated numerical parameters.

By storing parameter-generation and calculation functions as scripts in the database, the system can dynamically create a new, unique set of problems each time a student takes the test. This not only prevents students from memorizing fixed question sets but also promotes deeper understanding and problem-solving skills through repeated exposure to varied task scenarios.

Thus, each time the student passes the test, he receives a unique set of tasks generated just for him.

To increase the motivation of students in the programs, it is advisable to implement the game principle. The prototype uses playful, step-by-step interaction to keep students engaged. By "peeling the nuts" (solving each part of the problem in sequence) and assembling correct answers like puzzle pieces, learners gain a sense of exploration and achievement. Game elements such as immediate feedback, hints, and visual cues transform the learning process into an enjoyable experience.

Students have the option to either enter the final answer directly or work through each stage of the solution. This approach fosters incremental learning: students can pinpoint exactly where they go wrong and receive immediate guidance on how to correct it. Over time, these targeted corrections reinforce strong problem-solving habits.

By tracking each step and offering just-in-time assistance, the system helps students identify and rectify misconceptions before they become entrenched. This ultimately leads to more stable competencies in arithmetic, algebraic manipulations, and other targeted math skills.

In Fig. 10 shows screen forms of the process of step-by-step solution of one of the NMT problems. The student has the opportunity to immediately enter the answer, or to solve the problem step by step, peeling the “nuts”. At each step, he assembles the correct answer, like a puzzle, by dragging the pieces of the answer (correct or incorrect) into the corresponding fields. If necessary, he can get additional information.

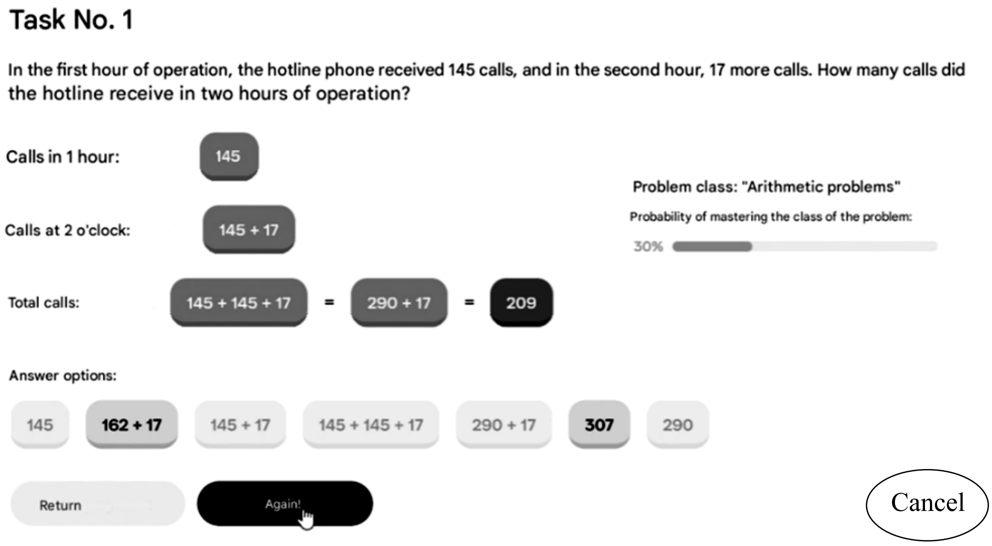


Fig. 10. An example of a screen for a step-by-step solution to a mathematical problem

In this way, the student in a playful way solves a number of tasks of a certain class, forming and consolidating the relevant competencies.

CONCLUSIONS

We have presented an author’s vision of a modern Intelligent Tutoring System (ITS) for exact sciences, particularly mathematics, aimed at preparing Ukrainian schoolchildren for the National Multi-Subject Test. A functional prototype has been developed to demonstrate how diagnostic control, real-time feedback, and adaptive guidance can be effectively combined.

Our prototype employs structured knowledge bases, algorithmic problem generation, and iterative error diagnosis. This design ensures that each step of a student’s solution can be systematically monitored and corrected, demonstrating a high level of scientific rigor in both pedagogy and software engineering.

Through an analysis of existing tools and research, we identified critical gaps in personalized learning and feedback mechanisms within mathematics education. The prototype addresses these gaps by incorporating specialized modules for real-time assessment and targeted remediation.

Based on the prototype’s initial success, the future goal is to refine and deploy the ICNP software throughout Ukraine. Further development will focus on expanding the system's capabilities, integrating additional educational resources, and conducting large-scale trials to validate its effectiveness and scalability.

REFERENCES

1. S. Van Vaerenbergh, A. Pérez-Suay, "A Classification of Artificial Intelligence Systems for Mathematics Education," in *Mathematics Education in the Age of Artificial Intelligence*, P.R. Richard, M.P. Vélez, and S. Van Vaerenbergh, Eds., vol. 17, Springer, Cham, 2022. doi: https://doi.org/10.1007/978-3-030-86909-0_5
2. N.N. AbuEloun, "Intelligent tutoring system for mathematics," *International Journal of Advanced Scientific Research*, vol. 2, no. 1, pp. 11–16, 2017.
3. S.-C. Shih, C.-C. Chang, B.-C. Kuo, and Y.-H. Huang, "Mathematics intelligent tutoring system for learning multiplication and division of fractions based on diagnostic teaching," *Educ. Inf. Technol.*, vol. 28, pp. 9189–9210, 2023. doi: https://doi.org/10.1007/s10639-022-11553-z
4. G. Wang, X. Chen, D. Zhang, Y. Kang, F. Wang, and M. Su, "Development and application of an intelligent assessment system for mathematics learning strategy among high school students—Take Jianzha County as an example," *Sustainability*, vol. 14, no. 19, p. 12265, 2022. doi: https://doi.org/10.3390/su141912265
5. A.S. Kulik, A.G. Chukhray, and O.V. Havrylenko, "Information technology for creating intelligent computer programs for training in algorithmic tasks. Part 1: Mathematical foundations," *System Research and Information Technologies*, no. 4, pp. 27–41, 2022. doi: https://doi.org/10.20535/SRIT.2308-8893.2021.4.02
6. A. Kulik, A. Chukhray, M. Chernenko, and O. Leshchenko, "Diagnostic support of an intelligent tutor system for teaching skills to solve algebraic equations," *Radioelectronic and Computer Systems*, no. 3(99), pp. 160–168, 2021. doi: https://doi.org/10.32620/reks.2021.3.13
7. A. Chukhray, E. Yashina, and O. Leshchenko, "Formalization of tasks generation for complex of interactive web-tests on math," *Radioelectronic and Computer Systems*, no. 3 (99), pp. 140–148, 2021. doi: https://doi.org/10.32620/reks.2021.3.11
8. *OnlineMSchool*. Available: https://ua.onlinemschool.com/math/assistance
9. *IXL*. Available: https://www.ixl.com/
10. *Khan Academy*. Available: https://khanacademy.org/
11. *LearnBop*. Available: https://www.learnbop.com/
12. *Mangahigh*. Available: https://www.mangahigh.com/en/
13. *GeoGebra*. Available: https://www.geogebra.org/
14. *CueThink*. Available: https://www.cuethink.com/
15. *WeBWorK*. Available: https://webwork.maa.org/wiki/WeBWorK_Main_Page
16. *An Internet Mathematics Assessment System*. Available: https://www.imathas.com/
17. S. Wolfram, "Wolfram|Alpha as the way to bring computational knowledge superpowers to ChatGPT," *Stephen Wolfram Writings*, Jan. 2023. Available: https://writings.stephenwolfram.com/2023/01/wolframalpha-as-the-way-to-bring-computational-knowledge-superpowers-to-chatgpt/
18. T. Trinh, T. Luong, "AlphaGeometry: An Olympiad-level AI system for geometry," *DeepMind Blog*, 2023. Available: https://deepmind.google/discover/blog/ alphageometry-an-olympiad-level-ai-system-for-geometry/
19. "Program for external independent assessment of learning outcomes in mathematics acquired on the basis of complete general secondary education," *Ministry of Education and Science of Ukraine*, Dec. 2019. Available: https://testportal.gov.ua/progmath/
20. "Official report on the results of the NMT in 2023," *Ukrainian Center for Educational Quality Assessment*, 2023. Available: https://testportal.gov.ua//wp-content/ uploads/2023/08/ZVIT-NMT_2023-Tom_1_.pdf

21. "NMT in mathematics. Demo version," *Ukrainian Center for Educational Quality Assessment*, 2024. Available: https://testportal.gov.ua/wp-content/uploads/2024/02/NMT_2024-Matematyka-Demo_sajt.pdf
22. A. Kulik, A. Chukhray, "The Principle of Control by Diagnosis – the Fundamental Principle of Computer-based Learning Process Intellectualisation," *2023 13th International Conference on Dependable Systems, Services and Technologies* (*DESSERT*)*, Athens, Greece, 2023*, pp. 1–6. doi: 10.1109/DESSERT61349.2023.10416498

Received 24.09.2024

INFORMATION ON THE ARTICLE

Anatoliy S. Kulik, ORCID: 0000-0001-8253-8784, National Aerospace University "Kharkiv Aviation Institute", Ukraine, e-mail: a.kulik@khai.edu

Oleg P. Zeleniak, ORCID: 0000-0003-0024-485X, Olexandriya Multi-Profile Lyceum, Ukraine, e-mail: zlnk@ukr.net

Andriy G. Chukhray, ORCID: 0000-0002-8075-3664, National Aerospace University "Kharkiv Aviation Institute", Ukraine, e-mail: a.chukhray@khai.edu

Oleksandr V. Prokhorov, ORCID: 0000-0003-4680-4082, National Aerospace University "Kharkiv Aviation Institute", Ukraine, e-mail: o.prokhorov@khai.edu

Olena S. Yashyna, ORCID: 0000-0003-2459-1151, National Aerospace University "Kharkiv Aviation Institute", Ukraine, e-mail: o.yashina@khai.edu

Olena V. Havrylenko, ORCID: 0000-0001-5227-9742, National Aerospace University "Kharkiv Aviation Institute", Ukraine, e-mail: o.havrylenko@khai.edu

Oleksandr O. Yevdokymov, ORCID: 0009-0008-9687-6344, National Aerospace University "Kharkiv Aviation Institute", Ukraine, e-mail: o.yevdokymov@khai.edu

Andrii A. Torzhkov, ORCID: 0009-0005-0257-095X, "SoftServe", Ukraine, e-mail: andrejtorzhkov94@gmail.com

Oleksii V. Zayarnyi, ORCID: 0009-0003-5208-3848, National Aerospace University "Kharkiv Aviation Institute", Ukraine, e-mail: o.zaiarnyi@khai.edu

КОНЦЕПЦІЯ ІНТЕЛЕКТУАЛЬНОЇ КОМП'ЮТЕРНОЇ НАВЧАЛЬНОЇ СИСТЕМИ ДЛЯ ПІДГОТОВКИ УКРАЇНСЬКИХ ШКОЛЯРІВ ДО НМТ / А.С. Кулік, О.П. Зеленяк, А.Г. Чухрай, О.В. Прохоров, О.С. Яшина, О.В. Гавриленко, О.О. Євдокимов, А.А. Торжков, О.В. Заярний

Анотація. Уже кілька років підготовка та проведення національного мультипредметного тесту (НМТ) в Україні відбувається в умовах онлайн-навчання. Результати тестування показують зниження успішності школярів у математиці. Подано прототип інтелектуальної навчальної системи для розв'язування математичних задач, яка має стати доступним засобом підготовки до тестування. Система забезпечує вирішення широкого кола математичних задач у покроковому режимі. Систему розроблено відповідно до принципу раціонального керування за діагнозом, що передбачає наявність багатьох діагностичних моделей. Це дає змогу поглиблено діагностувати помилки учнів. Інструменти штучного інтелекту дозволять реалізувати індивідуальні рекомендації для кожного учня з урахуванням його рівня підготовки і цілей навчання.

Ключові слова: інтелектуальна репетиторська система, національний мультипредметний тест, задачі з математики.

UDC 004.51
DOI: 10.20535/SRIT.2308-8893.2025.2.10

ANALYSIS OF WEB ACCESSIBILITY OF UKRAINIAN HIGHER EDUCATION INSTITUTIONS' WEBSITES

B.O. KUZIKOV, P.O. TYTOV, O.A. SHOVKOPLIAS

Abstract. In today's digital world, website content accessibility for all users, including people with disabilities, is crucial. This paper examines the accessibility of web content on the websites of Ukrainian higher education institutions to assess their compliance with modern standards and requirements. The goal is to identify problematic aspects and develop strategies for improving the accessibility of educational web resources. The study covered data from the main pages of 372 higher education institution websites, whose addresses were obtained from the Unified State Electronic Database on Education ("USEDE"). The built-in accessibility enhancement tools used on these pages were analyzed. The paper also summarizes global experience in regulating web accessibility requirements. The analyzed regulatory documents include Web Content Accessibility Guidelines version 2.1 in their rules. Automated analysis tools WAVE and Axe were used to assess website compliance with the Guidelines. Results showed that the most common problems were insufficient highlighting of hyperlinks (characteristic of 84.8% of sites, Success criteria (SC) 2.4.4, 4.1.2), insufficient element contrast (84.6%, SC 1.4.3), lack of alternative text for images (42.9%; SC 1.1.1), and non-compliance with markup and semantics requirements. The conclusions present an analysis of the problems identified during the study, along with the authors' recommendations for addressing them. The analysis results were systematized and published on a specially created web resource. The authors plan to conduct similar studies on an ongoing basis. It is also necessary to pay attention to problems identified when using automatic analysis tools. The Axe tool identified 11.875 cases of element contrast issues, accounting for 24% of the total number of detected problems. However, this figure is not final, as it does not account for the full range of possible color perception violations and overlooks the contrast of elements against backgrounds with non-uniform colors.

Keywords: web accessibility, WCAG, accessible education, inclusiveness, webpage.

INTRODUCTION

The shift of a significant portion of daily activities online has become one of the defining trends of modern times. Digitalization has encompassed almost all aspects of human life, including using electronic services to obtain certificates through the Diia app, scheduling doctor appointments through the Helsi information system, submitting gas, water, and electricity meter readings, purchasing tickets, online banking, etc. These are just some examples that illustrate the depth of electronic services' integration into our lives. Parallel to this, the importance of these services' accessibility to the broadest possible range of users is increasing. World Health Organization statistics show that at least 2.2 billion people have vision impairments [1]. According to the European Union's official website, 100 million people in the EU live with various forms of disability [2], making them a significant portion of potential users of digital products and services. Ensuring the accessibility of these products and services not only complies with ethical standards of equality and inclusiveness but can also bring economic benefits to companies by expanding their customer base. In Ukraine, where the number of people with special needs is increasing due to military actions, the relevance of digital accessibility is growing even more.

The author's research focuses on analyzing information services provided by higher education institutions through their web platforms, emphasizing the accessibility of these services. In the context of ensuring equal access to education, web accessibility of educational resources becomes particularly important, as inclusiveness ensures that every student, regardless of their special needs, will have comfortable learning conditions. According to the definition provided in the digital accessibility manual [3], web accessibility means designing and developing websites, web tools, and technologies in a way that makes them usable by people with disabilities. Web accessibility is a specific case of the broader concept of "digital accessibility", although sometimes both terms are used synonymously. Thus, studying barriers in accessing educational institution websites is key to identifying and overcoming obstacles to digital inclusion. This work aims to analyze the web accessibility of Ukrainian higher education institutions' websites based on data obtained from the Register of Educational Activity Entities of the Unified State Electronic Database on Education [4].

In the context of educational resources, attention to web accessibility is driven by both general humanistic factors, such as guaranteeing accessible education — equal access to information for all who desire it, promoting an inclusive environment, and purely practical considerations — enhancing reputation and expanding the circle of potential students. It's also worth considering "staying ahead of the curve". Several countries, including the USA, Great Britain, and European Union states, have already implemented legislation requiring compliance with minimum web accessibility standards for specific categories of websites. Given Ukraine's ambitions for European integration, the need to harmonize national legislation with international experience in this field is anticipated.

In parallel with progress in the field of information and communication technologies, regulatory acts supporting the inclusion of people with disabilities are emerging in the legislative field. This encompasses both framework documents, exemplified by Article 9 of the United Nations Convention on the Rights of Persons with Disabilities [5], Article 26 of the EU Charter of Fundamental Rights [6], and more specific legislative implementation. It's worth noting the experience of the USA – Section 508 of the Rehabilitation Act [7], the EU – Web Accessibility Directive (2016/2102) [8], Great Britain – Equality Act 2010 [9], international standard ISO/IEC 40500:2012 "Information technology – W3C Web Content Accessibility Guidelines (WCAG) 2.0" [10], Ukraine – state standard for digital accessibility National Standard of Ukraine (DSTU) EN 301 549:2022 "Information Technologies. Accessibility requirements for ICT products and services" [11], which duplicated the European standard of the same name that includes Web Content Accessibility Guidelines version 2.1, which European legislation has been using for several years. Overall, all these documents include WCAG [12] in their list of rules. Therefore, when analyzing website accessibility, we will use the WCAG standard and tools to verify compliance.

CURRENT STATE

Web accessibility testing is a method for verifying the accessibility of digital content for all users, including those with physical limitations or cognitive impairments. Such testing aims to ensure content accessibility through alternative interaction methods, not limited to traditional methods such as using a mouse or touchpad. The main principles of web accessibility include [13]:

- perception: the interface and information should be presented in a way that is accessible to the user. For example, text to background on the page or in images should have a contrast ratio of at least 4.5:1, and images should include a text description (alt-text);
- operability: navigation must be accessible and controllable through the user interface. For example, the ability to navigate to any element using a keyboard;
- understandability: control elements must be understandable and standardized within the resource;
- robustness: the user interface must ensure content accessibility for all users. This can usually be achieved with syntactically and semantically correct HTML markup and compliance with other related web specifications.

Within the framework of this study, the primary attention was paid to checking websites for compliance with perception criteria, as ensuring interface accessibility for different user groups is a key requirement of web development. Creating high-contrast content, proper use of alternative text for images, and other measures allow for making websites more accessible to people with various levels of vision and those who use assistive technologies. This creates an inclusive internet space where users can access information and interact with web resources comfortably and effectively.

To evaluate accessibility, different strategies are required: automated accessibility validators, manual verification, and expert evaluation. Website or application accessibility can be automatically assessed using several online tools. These tools can complete assessments of various WCAG versions at certain compliance levels and check specific aspects of accessibility, such as color contrast or the presence of necessary element attributes. Most criteria can be checked automatically. For example, "All non-text content that is presented to the user has a text alternative that serves the equivalent purpose" (Success criterion (SC) 1.1.1). However, certain aspects require manual verification beyond automated tests, as they concern the completeness of WCAG guidelines compliance. The most common web accessibility evaluation tools include WAVE, Axe, TAW, and Web Accessibility Inspector [14].

Most accessibility testing tools perform various types of checks. For example, WAVE [15], developed by WebAIM, is a browser toolbar that allows analyzing accessibility without storing data on a server. This tool identifies the need for manual checks of ARIA (Accessible Rich Internet Applications) elements, considering descriptive and precise values, adequate use of states, roles, properties, as well as correct use of tab indices and active regions. The Axe tool [16], created by Deque Systems, is also designed to help website developers identify and resolve accessibility issues on their sites according to WCAG recommendations. TAW [17], developed by CTIC Centro Tecnologico, is an automated accessibility testing tool that meets WCAG 1.0 and 2.0 criteria. It identifies accessibility issues and provides targeted recommendations for resolving them. Accessibility Inspector [18] tests desktop application accessibility, pointing out elements with issues after entering a URL or file path. It is compatible with Windows and Mac OS X.

REVIEW OF GLOBAL EXPERIENCE IN WEB ACCESSIBILITY

The article's authors [19] analyzed web accessibility in higher education in Great Britain. Considering WCAG 2.0 requirements, they examined the accessibility of the home pages of 66 research universities included in the SCImago international ranking. Three automated web accessibility tools were used for this purpose: TAW, WAVE, and EIII Page Checker. In total, the study examined 120 research universities. It was found that the most common violations were related to the lack of text alternatives for non-text content, contrast errors, and the need to optimize web pages using modern technologies and tools. In particular, color contrast errors between text and background were detected on 74.2% of the examined university home pages.

In the work [20], the authors conducted a comparative analysis of Turkish university websites regarding accessibility, usability, performance, and readability. The URL data of websites for verification were collected from the Turkish Higher Education Council (THEC) website. The sample covered 186 universities, among which five were vocational schools and seven were inactive or offline. Thus, 179 Turkish universities were evaluated using AChecker for compliance with WCAG 2.0 criteria. The biggest problems were at conformance level A. Private university sites showed worse conformance results at AA and AAA levels than public universities. Of 110 public university sites, only 10 complied with all three conformance levels. Only four of 69 private university sites met levels A, AA, and AAA. The most common problem on the sites was the lack of appropriate alternative text: this error was found on 69% of public university sites and 55% of private university sites.

Several studies in Ukraine have examined website accessibility using various methods and tools. For instance, in the work [21], documentary analysis and web-page verification methods using the WAVE tool were employed. The authors investigated library web content accessibility and established that most web pages do not comply with WCAG 2.0 and WCAG 2.1 standards. It was found that no large-scale study of library content accessibility had been conducted in Ukraine. A review of 22 regional universal scientific library sites showed that none meets WCAG 2.0 requirements, and the lack of attention to optimizing domestic library web content has effectively resulted in information inaccessibility. Additionally, 177 national library sites were checked, revealing that most use tools to engage people with special needs.

The article [22] presents the results of a study on the accessibility of Ukrainian higher education institutions' websites according to WCAG 2.1 guidelines, using the WAVE tool to identify and systematize accessibility issues. The study analyzed 299 home pages of higher education institution sites. The author proposes dividing all errors into six types, considering the criticality of the problem as a coefficient of a generalized metric for ranking resources. The metric is based on the ratio of problematic elements to the total number of elements in the category. However, more detailed information is not provided. The study emphasizes that the level of compliance of the studied sites with accessibility standards, even at the basic level, is low.

DATA COLLECTION FOR ANALYSIS

This work aims to analyze and evaluate the accessibility level of Ukrainian higher education institutions' websites to determine their compliance with modern standards and web accessibility requirements. The main tasks include identifying factors affecting web resource accessibility for different target audiences, including people with disabilities, and analyzing the implementation of adaptive design and other technologies that facilitate website access.

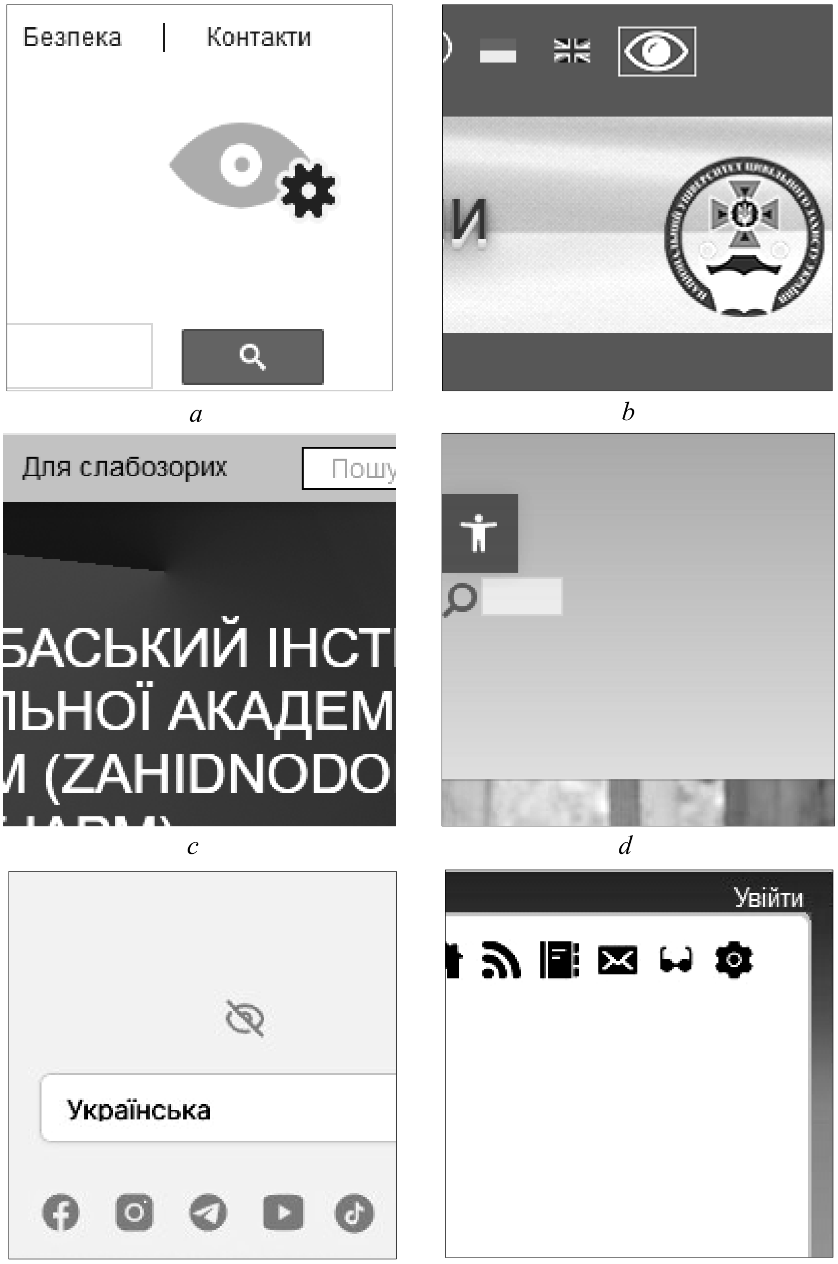
The list of higher education institutions was exported from the info.edbo.gov.ua website [4]. As of January 24, 2024, the list contained 571 entries, from which 134 entries of institutions that did not have website addresses specified or were categorized as revoked were removed. The address list was verified at the next stage using the regular expression `^((http|https)://)?[A-z0-9./]+$` revealed eight entries with errors such as typos, substituting Latin letters with Cyrillic ones, and others. The identified errors were corrected. After this, a crafted script checked the site availability via HTTP/HTTPS protocols with and without the www prefix. Preference was given to the version that responded with HTTP code 200 "OK". The new address was saved for sites that responded with codes 301/302/308 (temporary and permanent redirects). Such processing is necessary since not all sites in the registry have a specified protocol, and some sites are inaccessible via the protocol/prefix specified in the database or their combination. Inaccessible sites (domain name not found — 16 cases, timeout exceeded — 3 cases) or duplicated were filtered out. In 8 cases, sites were specified with HTTPS protocol and "www" prefix, but SSL certificates for these domain names were not created; these errors were corrected. After processing, 385 addresses remained on the list. However, even after processing, the data quality remains insufficient because some sites displayed server configuration error pages, were empty, showed hosting pages, or informed about domain registration expiration. Content analysis also revealed that some sites listed in the Registry function as e-shops and online casinos.

For sites that were accessible during the period from February 1, 2024, to February 2, 2024, a full-page screenshot was automatically taken in Firefox browser, the page's HTML code was saved for further analysis, and a list of text blocks was preserved, extended with coordinates on the page and HTML attributes. The Firefox browser was chosen due to its capability to take screenshots of the entire page. For all sites, the window width was 1349 pixels.

CHECKING FOR TOOLS FOR PEOPLE WITH VISUAL IMPAIRMENTS

Through visual analysis of website screenshots, those that contained easily identifiable adaptation tools for people with visual impairments in the visible part of the full-screen browser mode were selected. Among those analyzed, only 25 sites had such tools. For example, Fig. 1 shows 200×200 pixel page fragments at 100% scale containing special adaptation tools.

Among the examined tools for implementing webpage adaptation functionality, four cases revealed the use of the POJO plugin [23], three used VBI technology [24], and specialized extensions for WordPress, Joomla, and Drupal. In four cases, the sources and authorship of the script could not be established. In most cases (20 from 25) the tool allows changing font size and switching the page to black and white or high contrast mode; in 4 cases — only font size can be changed, and in one case (West Ukrainian National University) — the page image can only be converted to grayscale.



a *b* *c* *d* *e* *f*

Fig. 1. Typical designations of adaptation tools on the example of sites of the Open International University of Human Development "Ukraine" (*a*); National University of Civil Defense of Ukraine (*b*); West Donbas Institute of Interregional Academy of Personnel Management (*c*); Dnipropetrovsk State University of Internal Affairs (*d*); Mariupol State University (*e*); Kherson State University (*f*)

During the analysis, cases were also recorded where plugin use made the site unreadable due to text block overlapping (Ivan Franko National University of Lviv), an inactive button labeled "ДЛЯ ЛЮДЕЙ З ПОРУШЕННЯМ ЗОРУ"("FOR PEOPLE WITH VISUAL IMPAIRMENTS", Ivano-Frankivsk National Medical University), plugin configuration error (text "[bvi text = "Button visually impaired"]" at Zhytomyr Institute of Economics and Humanities — a separate structural unit of the Open International University of Human Development "Ukraine").

We positively evaluate cases of application or inclusion in the technical requirements for university website development of specialized tools, although their implementation was not correct in all cases. The capabilities provided by such extensions, such as font enlargement and contrast schemes, can be implemented using standard tools or browser extensions, allowing people to adapt them to their own needs. Using plugins is appropriate when access to the site repeatedly occurs from unprepared devices, such as computers in classrooms or public places. This scenario is more characteristic of the university's internal resources, such as distance learning systems, electronic cabinets, schedules, etc., which are not objects of this study. Considering the above, adapting the entire site to WCAG recommendations is more appropriate than applying separate extensions. Such verification was performed at the next stage.

DISCUSSION AND FURTHER RESEARCH DIRECTIONS

In today's digital world, websites must be accessible to all users, including people with disabilities. To assess the state of web accessibility of Ukrainian higher education institutions' main pages, WAVE and Axe were used as automated tools for checking WCAG compliance.

WAVE Web Accessibility Evaluation Tool

For web accessibility analysis, 377 home pages of Ukrainian higher education institutions were selected. Each page underwent verification using WAVE, and the obtained results were recorded in a table containing information about the educational institution's name, verification status, number of WCAG violations, number of contrast errors, and number of ARIA elements on the page. This analysis method provided a high degree of automation and objectivity in determining the web accessibility level of university pages, providing objective data for further scientific study and development of web accessibility improvement strategies.

When analyzing the results, we would like to draw special attention to three indicators: the total number of violations, the most common violation (insufficient element contrast), and the use of ARIA elements. The number of WCAG violations indicates the level of failure of basic web accessibility principles. The study found that the following higher education institutions' websites had the most WCAG failures: International European University — 322 violations, Sumy Applied College of Construction and Architecture — 286, Odesa National Economic University — 277. The number of contrast failures helps determine how convenient it is for people with various visual impairments to use the website. The following educational institutions had the most contrast failures: Chortkiv Educational and Scientific Institute of Entrepreneurship and Business WUNU — 212 violations, National Forestry University of Ukraine — 197, Mykolaiv National Agrarian University — 190. Assessment of the number of ARIA elements shows how effectively technologies are used to facilitate website perception for users with disabilities. The following websites had the most ARIA elements: Bila Tserkva Institute of Continuous Professional Education — 925, Lviv State University of Physical Culture named after Ivan Boberskyi — 864, Vasyl Stus Donetsk National University — 835.

Generalized data on the number of violations found on websites using WAVE are shown in Fig. 2. The graph is limited to 130 failures because larger values are isolated cases. The absolute number of failures in the indicator itself is not very informative because it does not consider their criticality and the total number of elements on the page. Moreover, the absolute number may be inaccurate since the tool can analyze specific categories of problems and may include false positive and false negative results in the report. However, any value greater than 0 indicates the presence of problematic elements on the site that need to be fixed. Many problems may suggest that web accessibility principles were not considered during site design. The distinction between "minor" and "significant" numbers of errors is subjective and depends on the nature of the errors and the site's content.

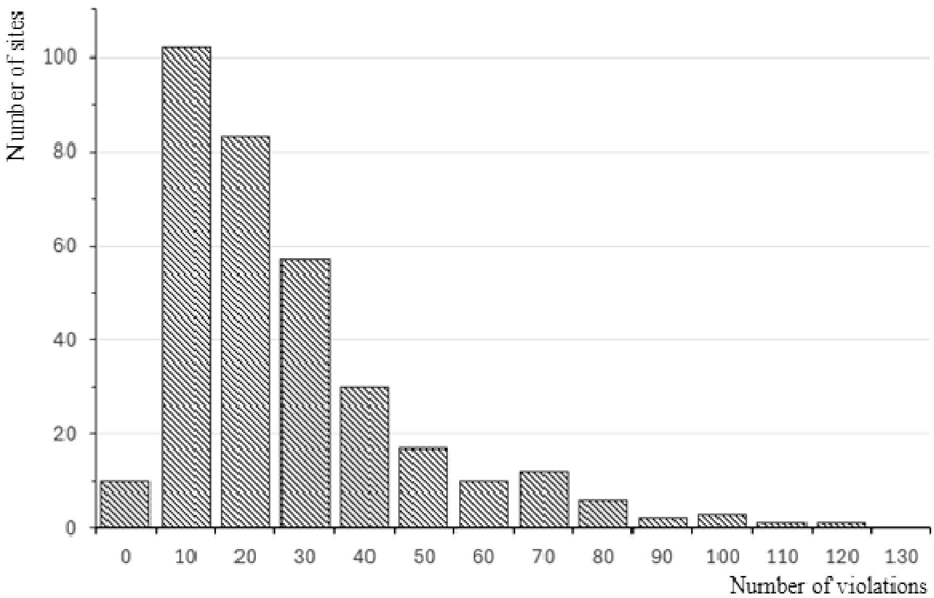


Fig. 2. Distribution of sites by number of accessibility errors (according to WAVE data)

Axe Web Accessibility Evaluation Tool

Using Axe allows automatic analysis of 15 out of 32 WCAG 2.1 criteria at compatibility level A and 4 out of 24 at level AA, reducing the response for each criterion to the categories "Passed," "Violation," and "Cannot be Applied" (for example, if the site does not contain elements of a specific category). The list of errors and remarks is generated as a separate report. Generalized statistics on typical problems and their prevalence on the studied sites are shown in Table, Fig. 3. The relative value (%) is the percentage of errors in the category based on the total number of errors or sites from the total number of sites in study. The analysis also considered rules classified as "Best Practice".

Statistics of typical web accessibility problems on HEI website pages based on Axe reports

Type of Violation	Number of Violations		Prevalence (number of sites)		WCAG Criterion
Links must be visually distinguished from surrounding text through styling or contrast.	23977	48.81%	319	84.84%	2.4.4, 4.1.2
Elements must meet minimum contrast ratio requirements	11875	24.17%	318	84.57%	1.4.3
Images must provide alternative text	7865	16.01%	159	42.29%	1.1.1
Buttons must have descriptive text	1295	2.64%	78	20.74%	4.1.2
Frames must have descriptive titles	528	1.07%	82	21.81%	4.1.2
Form elements must have associated labels	520	1.06%	39	10.37%	4.1.2
Links must be distinguishable by means other than color	502	1.02%	92	24.47%	1.4.1
Elements must only use valid ARIA attributes	390	0.79%	76	20.21%	4.1.2
Hidden ARIA elements must not be focusable nor contain focusable elements	311	0.63%	31	8.24%	4.1.2
List items ( ) must be contained within or parent elements	224	0.46%	24	6.38%	1.3.1

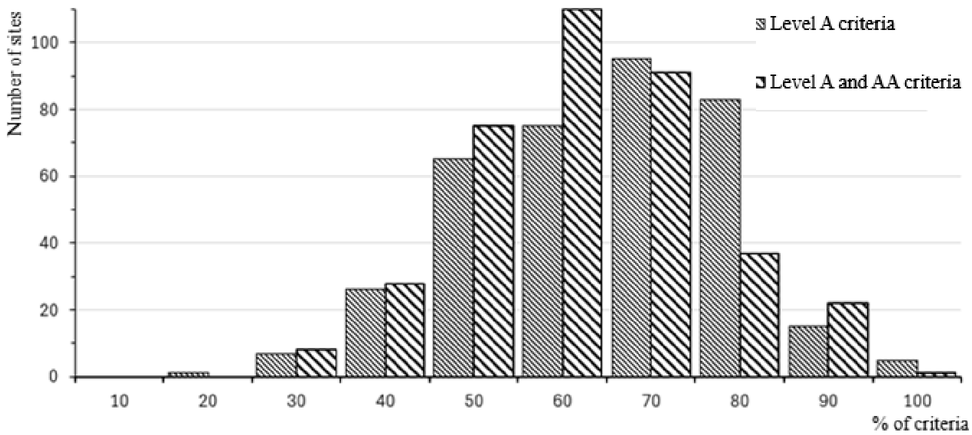


Fig. 3. Distribution of sites by percentage of WCAG criteria compliance, Axe analysis

Based on the results of the analysis, further research directions were identified.

1. Standard promotion and raising awareness about web accessibility criteria are essential. The Web Accessibility HUB website [25] created by us already contains the results presented in the article and will be updated with future analytics.

2. Improvement of automatic website verification tools. Analysis of reports revealed the following problems:

- Contrast issues rank second in prevalence (24.17% of all detected problems found on 84.5% of sites). Existing tools have deficiencies in determining the contrast ratio for text that is part of an image or placed on a non-uniform background. Also, automated tools do not investigate the relationship between color perception and object contrast.

- Reports obtained by automated verification tools contained Type I and Type II errors, mainly when checking the semantics of ARIA labels and role attribute settings [26]. Based on the collected information, we created a dataset on

Kaggle [27], which will be used to develop tools based on artificial intelligence and machine vision. This will improve the quality of automatic detection of such problems.

3. Development of synthetic pages to identify deficiencies in automatic failures detection tools. As part of this research, specific test pages and additional rules for axe-core have been developed. These findings will be published in the paper "Detection and Prevention of Accessibility Cloaking Attacks", which is currently being prepared for publication.

CONCLUSION

Effective implementation of web resource accessibility in higher education institutions requires a comprehensive approach in the context of accessible education, which includes technical aspects, legislative support, as well as educational activities. Along with improving the technical and legislative aspects of web accessibility, it is necessary to focus on increasing society's awareness of this issue. Spreading information about the importance of web accessibility and its significance for different user groups can draw more attention to the problem and strengthen support for necessary measures.

The study of web accessibility of Ukrainian higher education institutions' websites revealed some significant problems concerning non-compliance with basic WCAG standards and criteria. Analysis using WAVE and Axe tools showed that the most common violations are contrast problems, lack of alternative texts for images, and incorrect use of ARIA labels. Such identified violations indicate the need to improve web resources to adapt them for users with various visual impairments and ensure general information accessibility.

The analysis results were systematized and published on the authors' specially created web resource.

REFERENCES

1. J.D. Steinmetz et al., "Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: the Right to Sight: an analysis for the Global Burden of Disease Study," *Lancet Glob. Health*, vol. 9, no. 2, pp. e144-e160, Feb. 2021. doi: 10.1016/S2214-109X(20)30489-7
2. "Web Accessibility," *European Commission. Shaping Europe's digital future*. Accessed on: Jan. 7, 2025. [Online]. Available: https://digital-strategy.ec.europa.eu/en/policies/web-accessibility
3. D. Popov, "Brief Digital Accessibility Manual," *Diia. Digital Education*. Kyiv, 2023. [Online]. Available: https://www.undp.org/uk/ukraine/publications/brief-guide-on-digital-accessibility
4. *Higher Education Institutions | Register of Educational Activity Subjects*. [Online]. Available: https://registry.edbo.gov.ua/
5. "Convention on the Rights of Persons with Disabilities - Manual for Human Rights Education with Young people," *Council of Europe*. [Online]. Available: https://www.coe.int/en/web/compass/convention-on-the-rights-of-persons-with-disabilities
6. K. Chupina, *Youth and Disabilities EU and CoE policy framework*. [Online]. Available: https://pjp-eu.coe.int/documents/42128013/47261683/Youth_and_Disabilities_CoE-EC_policy_framework_KC_final.pdf/1ae78341-b755-4b73-9c0a-eda22cc7f08f
7. "Section 508 of the Rehabilitation Act of 1973," *U.S. Congress*. [Online]. Available: https://www.section508.gov/manage/laws-and-policies/

8. "Directive (EU) 2016/2102 on the accessibility of the websites and mobile applications of public sector bodies," *The European Parliament and the Council of the European Union*, Oct. 26, 2016. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016L2102
9. "Equality Act 2010," *legislation.gov.uk.* [Online]. Available: https://www.legislation.gov.uk/ukpga/2010/15/contents
10. "Information technology — W3C Web Content Accessibility Guidelines (WCAG) 2.0." *ISO/IEC 40500:2012*. [Online]. Available: https://www.iso.org/standard/ 58625.html
11. "Ukrainian developers will be able to create more online resources accessible for people with visual or hearing impairments. International WCAG 2.1 guidelines translated into Ukrainian for the first time," *Ministry of Digital Transformation of Ukraine*. [Online]. Available: https://thedigital.gov.ua/news/ukrainski-rozrobniki-zmozhut-stvoryuvati-bilshe-onlayn-resursiv-dostupnikh-dlya-lyudey-z-porushennyam-zoru-chi-slukhu
12. "Web Content Accessibility Guidelines (WCAG) 2.1." *W3C*. [Online]. Available: https://www.w3.org/TR/WCAG21/
13. "Accessibility Principles," *State Websites Design System*. [Online]. Available: https://design.gov.ua/ua/principy-dostupnosti
14. "11 Best Website Accessibility Testing Tools for 2022," *UI-licious*. [Online]. Available: https://uilicious.com/blog/best-website-accessibility-testing-tools/
15. "Web Accessibility Evaluation Tools – WAVE," *WebAIM* [Online]. Available: https://wave.webaim.org/
16. "axe: Accessibility Testing Tools and Software," *Deque Systems*. [Online]. Available: https://www.deque.com/axe/
17. "Web accessibility and W3C standardization services," *TAW*. [Online]. Available: https://www.tawdis.net/
18. "Accessibility Inspector," *Mozilla Firefox Source Docs documentation*. [Online]. Available: https://firefox-source-docs.mozilla.org/devtools-user/accessibility_inspector/
19. S. Alim, "Web accessibility of the top research-intensive universities in the UK," *SAGE Open*, vol. 11, no. 4, pp. 1–9, Oct. 2021. doi: 10.1177/21582440211056614
20. Y. Akgül, "Accessibility, usability, quality performance, and readability evaluation of university websites of Turkey: A comparative study of state and private universities," *Universal Access Inf. Soc.*, Apr. 2020. doi: 10.1007/s10209-020-00715-w
21. O.Y. Marina, S.O. Marin, "Versions of WCAG and accessibility of library web-content," *Visnyk of Kharkiv State Academy of Culture*, no. 56, Dec. 2019. doi: 10.31516/2410-5333.056.08
22. R.S. Savitskyi, "Accessibility of web portals of educational institutions of Ukraine," *Tech. Engineering*, no. 1(91), pp. 172–177, Jul. 2023. http://doi:10.26642/ten-2023-1(91)-172-177
23. "One click accessibility," *Pojo.me*. Accessed on: Jan. 7, 2025. [Online]. Available: https://pojo.me/plugins/accessibility
24. "Veks/buttonButton visually impaired," *GitHub*. Accessed on: Jan. 7, 2025. [Online]. Available: https://github.com/veks/button-visually-impaired-javascript
25. B. Kuzikov, *Web Accessibility HUB*. Accessed: Jan. 7, 2025. [Online]. Available: https://web-accessibility.sumdu.edu.ua
26. P. Tytov, B. Kuzikov, "Formation of the training sample of the information system for assessing the compliance of sites with web accessibility requirements," in *Proc. Informatics, Mathematics, Automation (IMA-2024)*, *Sumy, Ukraine, Apr. 26, 2024*, pp. 172–173.
27. P. Tytov, "Aria elements," *Kaggle*. [Online]. Available: https://www.kaggle.com/datasets/tytovpavel/aria-elements

Received 12.04.2024

INFORMATION ON THE ARTICLE

Borys O. Kuzikov, ORCID: 0000-0002-9511-5665 , Sumy State University, Ukraine, email: b.kuzikov@dl.sumdu.edu.ua

Pavlo O. Tytov, ORCID: 0009-0003-6911-5463, Sumy State University, Ukraine, email: stegaspasha@gmail.com

Oksana A. Shovkoplias, ORCID: 0000-0002-4596-2524, Sumy State University, Ukraine, email: o.shovkoplyas@mss.sumdu.edu.ua

АНАЛІЗ ВЕБДОСТУПНОСТІ САЙТІВ УКРАЇНСЬКИХ ЗАКЛАДІВ ВИЩОЇ ОСВІТИ / Б.О. Кузіков, П.О. Титов, О.А. Шовкопляс

Анотація. У сучасному цифровому світі доступність контенту веб-сайтів для всіх користувачів, у тому числі людей з обмеженими можливостями, є надзвичайно важливою. У роботі досліджується доступність вебконтенту на сайтах закладів вищої освіти України з метою оцінки їх відповідності сучасним стандартам та вимогам. Кінцевою метою є виявлення проблемних аспектів та розроблення стратегій для покращення доступності освітніх вебресурсів. Дослідження охопило дані щодо головних сторінок 372 сайтів закладів вищої освіти, адреси яких отримано із Єдиної державної електронної бази з питань освіти. Проаналізовано вбудовані засоби підвищення доступності, використані на цих сторінках. Узагальнено світовий досвід регулювання вимог до вебдоступності. Проаналізовано нормативні документи, які включають у перелік своїх правил Настанови з доступності вебвмісту версії 2.1. Для оцінювання відповідності вебсайтів Настановам використано автоматизовані інструменти аналізу WAVE та Axe. Результати показали, що найбільш поширеними проблемами були: недостатнє виділення гіперпосилань (характерно для 84,8% сайтів, критерій 2.4.4, 4.1.2 Настанов), недостатній контраст елементів (84,6%, критерій 1.4.3 Настанов), відсутність альтернативного тексту для зображень (42,9%; критерій 1.1.1 Настанов), недотримання вимог до розмітки та семантики. Висновки містять аналіз проблем, виявлених у ході дослідження, та рекомендації авторів щодо методів їх вирішення. Результати аналізу систематизовано й опубліковано на спеціально створеному вебресурсі. Автори планують здійснювати подібні дослідження на постійній основі. Необхідно звернути увагу на проблеми, виявлені при використанні інструментів автоматичного аналізу. Інструмент Axe ідентифікував 11875 випадків проблем із контрастністю елементів, що становить 24% від загальної кількості виявлених проблем. Проте ця цифра не є остаточною, оскільки не враховується повне коло можливих порушень кольоросприняття та ігнорується контрастність елементів на фоні з нерівномірним кольором.

Ключові слова: вебдоступність, WCAG, доступна освіта, інклюзивність, вебсторінка.

ВІДОМОСТІ ПРО АВТОРІВ

Бизов Іван Сергійович,
аспірант кафедри математичного моделювання та аналізу даних Навчально-наукового інституту комп'ютерних наук та штучного інтелекту Харківського національного університету імені В.Н. Каразіна, Україна, Харків

Бологін Андрій Сергійович,
старший науковий співробітник, кандидат технічних наук, начальник науково-дослідного управління експлуатації військової авіаційної техніки Державного науково-дослідного інституту авіації, Україна, Київ

Бологіна Юлія Олександрівна,
науковий співробітник Державного науково-дослідного інституту авіації, Україна, Київ

Бондаренко Валерія Вікторівна,
кандидат технічних наук, дослідник Прибережного університету, Франція, Кале

Бондаренко Віктор Григорович,
професор, доктор фізико-математичних наук, професор кафедри математичних методів системного аналізу ННІПСА КПІ ім. Ігоря Сікорського, Україна, Київ

Вещиков Георгій Вячеславович,
аспірант кафедри теоретичної електротехніки факультету електроенерготехніки та автоматики КПІ ім. Ігоря Сікорського, Україна, Київ

Гавриленко Олена Валеріївна,
доцент, кандидат фізико-математичних наук, доцент кафедри інформаційних систем і технологій факультету інформатики та обчислювальної техніки КПІ ім. Ігоря Сікорського, Україна, Київ

Гавриленко Олена Володимирівна,
доцент, кандидат технічних наук, доцент кафедри систем управління літальними апаратами Національного аерокосмічного університету "Харківський авіаційний інститут", Україна, Харків

Данилов Валерій Якович,
професор, доктор технічних наук, професор кафедри математичних методів системного аналізу ННІПСА КПІ ім. Ігоря Сікорського, Україна, Київ

Даць Ірина Вільямівна,
народна артистка України, професор кафедри оперної педагогіки та музичної режисури Національної музичної академії України імені П. І. Чайковського, Україна, Київ

Євдокимов Олександр Олегович,
магістр, аспірант кафедри математичного моделювання та штучного інтелекту Національного аерокосмічного університету "Харківський авіаційний інститут", Україна, Харків

Заярний Олексій Володимирович,
магістр, аспірант кафедри математичного моделювання та штучного інтелекту Національного аерокосмічного університету "Харківський авіаційний інститут", Україна, Харків

Зеленяк Олег Петрович,
Заслужений вчитель України, кандидат педагогічних наук, вчитель вищої категорії Олександрійського мультипрофільного ліцею, Україна, Олександрія

Кузіков Борис Олегович,
доцент, кандидат технічних наук, доцент кафедри комп'ютерних наук Сумського державного університету, Україна, Суми

Кулік Анатолій Степанович,
професор, доктор технічних наук, професор кафедри систем управління літальними апаратами Національного аерокосмічного університету "Харківський авіаційний інститут", Україна, Харків

Маслянко Павло Павлович,
старший науковий співробітник, кандидат технічний наук, доцент кафедри прикладної математики факультету прикладної математики КПІ ім. Ігоря Сікорського, Україна, Київ

Нікітін Владислав Олегович,

аспірант ННІПСА КПІ ім. Ігоря Сікорського, Україна, Київ

Островерхов Микола Якович,

професор, доктор технічних наук, завідувач кафедри теоретичної електротехніки факультету електроенерготехніки та автоматики КПІ ім. Ігоря Сікорського, Україна, Київ

Петренко Анатолій Іванович,

професор, доктор технічних наук, професор кафедри системного проектування ННІПСА КПІ ім. Ігоря Сікорського, Україна, Київ

Прохоров Олександр Валерійович,

професор, доктор технічних наук, професор кафедри комп'ютерних наук та інформаційних технологій Національного аерокосмічного університету "Харківський авіаційний інститут", Україна, Харків

Романов Микола Дмитрович,

студент магістратури факультету прикладної математики КПІ ім. Ігоря Сікорського, Україна, Київ

Сільвестров Антон Миколайович,

професор, доктор технічних наук, професор кафедри теоретичної електротехніки факультету електроенерготехніки та автоматики КПІ ім. Ігоря Сікорського, Україна, Київ

Спінул Людмила Юріївна,

доцент, кандидат технічних наук, доцент кафедри теоретичної електротехніки факультету електроенерготехніки та автоматики КПІ ім. Ігоря Сікорського, Україна, Київ

Тимчук Юрій Михайлович,

заступник начальника 4993 Військового представництва Міністерства оборони України, Україна, Одеса

Титов Павло Олегович,

аспірант кафедри комп'ютерних наук Сумського державного університету, Україна, Суми

Торжков Андрій Андрійович,

магістр, Senior Java Software Engineer "SoftServe"

Фещенко Кирил Юрійович,

аспірант кафедри інформаційних систем і технологій факультету інформатики та обчислювальної техніки КПІ ім. Ігоря Сікорського, Україна, Київ

Халімовський Олексій Модестович,

доцент, кандидат технічних наук, доцент кафедри теоретичної електротехніки факультету електроенерготехніки та автоматики КПІ ім. Ігоря Сікорського, Україна, Київ

Чухрай Андрій Григорович,

професор, доктор технічних наук, професор кафедри інженерії програмного забезпечення Національного аерокосмічного університету "Харківський авіаційний інститут", Україна, Харків

Шовкопляс Оксана Анатоліївна,

доцент, кандидат фізико-математичних наук, завідувач кафедри комп'ютерних наук Сумського державного університету, Україна, Суми

Яковлев Сергій Всеволодович,

член-кореспондент НАН України, доктор фізико-математичних наук, професор, Заслужений діяч науки та техніки України, Лауреат державної премії України, заступник директора з наукової роботи Навчально-наукового інституту комп'ютерних наук та штучного інтелекту Харківського національного університету імені В.Н. Каразіна, Україна, Харків

Яшина Олена Сергіївна,

доцент, кандидат технічних наук, доцент кафедри комп'ютерних наук та інформаційних технологій Національного аерокосмічного університету "Харківський авіаційний інститут", Україна, Харків