

МІНІСТЕРСТВО ОСВІТИ ТА НАУКИ УКРАЇНИ  
ДЕРЖАВНИЙ ТОРГОВЕЛЬНО-ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ  
ДЕРЖАВНИЙ ТОРГОВЕЛЬНО-ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ

Кваліфікаційна наукова  
праця на правах рукопису

**ХОРОЛЬСЬКА КАРИНА ВІКТОРІВНА**

УДК 004.92:004.032.26

**ДИСЕРТАЦІЯ**  
**ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ РОЗПІЗНАВАННЯ ГРАФІЧНОЇ ІНФОРМАЦІЇ**  
**НА ОСНОВІ НЕЙРОННОЇ МЕРЕЖІ**

122 – Комп'ютерні науки  
(шифр і назва спеціальності)

Інформаційні технології  
(галузь знань)

Подається на здобуття наукового ступеня доктор філософії

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

\_\_\_\_\_ К. В. Хорольська

Науковий керівник Харченко Олександр Анатолійович, кандидат технічних наук, доцент

**Київ-2023**

## АНОТАЦІЯ

Хорольська К.В. Інформаційна технологія розпізнавання графічної інформації на основі нейронної мережі. – Кваліфікаційна робота на правах рукопису.

Дисертація на здобуття наукового ступеня доктор філософії в галузі знань «Інформаційні технології» за спеціальністю 122 «Комп'ютерні науки». Державний торговельно-економічний університет, Київ, 2023.

Дисертація є комплексним дослідженням з моделювання, розробки та застосування інформаційної технології розпізнавання графічної інформації на основі нейронної мережі.

Актуальність теми дослідження обумовлюється факторами важливості розпізнавання графічної інформації, практичного застосування теоретичних основ розпізнавання графічної інформації на основі високоефективних рішень. На даному етапі розвитку практичного застосування теоретичних основ розпізнавання графічної інформації залишаються невирішеними проблеми: якісного виділення контурів; автоматизованого виділення об'єктів; варіабельності просторового розташування об'єктів та їх розпізнавання; практичного застосування методологічних розробок в розпізнаванні графічної інформації; якісної класифікації об'єктів розпізнавання.

Розпізнавання візуальних образів є суттєвим компонентом автоматизованих та/або інтелектуальних систем управління та обробки інформації. Задачі, пов'язані з ідентифікацією і класифікацією об'єктів та сигналів на основі обмеженого набору властивостей, є актуальними в галузях, таких як робототехніка, інформаційний пошук, моніторинг і аналіз візуальних даних, а також у дослідженнях штучного інтелекту. Алгоритмічна обробка та класифікація зображень застосовуються в системах безпеки, контролю доступу, віртуальній реальності та інформаційних пошукових системах. У зв'язку з розширенням

використання систем віртуальної реальності та розвитком гейм індустрії, а також у зв'язку з тим, що створення тривимірних моделей передбачає попереднє перетворення двовимірних зображень, виникає питання про необхідність ефективного процесу перетворення 2D зображень у 3D моделі.

Чималу роль у вирішенні зазначених проблем набуває створення програмних комплексів і математичного апарату, заснованих на застосуванні нейронних мереж, експертних систем і кібернетики.

Особлива увага приділена штучним нейронним мережам (або просто нейронним мережам) – обчислювальним системам, що вчаться на даних спостережень за допомогою процесу оптимізації, коли параметри моделі ітеративно коригуються, щоб мінімізувати різницю між прогнозованим результатом і фактичним результатом. Нейронні мережі, і, зокрема, згорткові нейронні мережі (CNN- convolutional neural network), широко визнані своєю ефективністю в аналізі візуальних зображень. CNN – це клас моделей глибокого навчання, спеціально обладнаних для моделей глибокого навчання, спеціально обладнані для обробки 2D-зображень. Вони складаються з кількох шарів невеликих нейронних колекцій, що обробляють частини вхідного зображення, які називаються рецептивними полями. Результати цих колекцій потім розміщуються так, щоб вони накладалися один на одного для кращого представлення вихідного зображення, що є ключовою особливістю CNN. Крім того, вони інваріантні до перекладу, тобто вони мають здатність ідентифікувати об'єкт як однаковий, коли він з'являється в різних поданнях. Ці функції надають CNN здатність фіксувати складні шаблони в просторовій і часовій областях - аспект, який є ключовим у завданні 3D-реконструкції з 2D-зображень. Штучні нейронні мережі також мають перевагу в обробці зашумлених, неповних або неоднозначних даних - сценарій, який зазвичай зустрічається в задачах обробки зображень. Вони здатні витягувати значущі характеристики навіть із недосконалих даних (зашумлення, пропуски, дублікати, протиріччя), забезпечуючи таким чином надійність роботи

моделі. На додаток до цього, нейронні мережі, а точніше CNN, мають здатність розпізнавати ієрархічні шаблони в даних. Наприклад, у завданнях обробки зображень вони можуть ідентифікувати краї та градієнти кольорів на нижчому рівні, а форми чи частини об'єкта – на вищому рівні. Ця функціональність має першочергове значення в таких завданнях, як 3D-реконструкція, які вимагають, щоб модель виявляла високорівневі особливості та зв'язки у 2D-зображеннях.

Основна увага в дисертаційному дослідженні приділена розпізнанню графічної інформації на основі нейронної мережі, тобто обробці графічних зображень у вигляді креслень та їх подальшої трансформації у 3D-модель з використанням логічних процедур навчання (ЛПН) згорткової нейронної мережі (CNN).

Інформаційною базою дослідження стали нормативні матеріали, наукові праці сучасних вітчизняних та закордонних науковців і практиків та емпіричні результати власних досліджень.

Теоретичною основою роботи є модель класифікатора графічних зображень на основі покриттів класів та елементарних класифікаторів примітивів для підвищення ефективності навчання CNN, а також принципи конструювання логічних процедур навчання з використанням апарату логічних функцій, що дозволить на практиці створювати ефективні програмні рішення CNN у задачах розпізнавання креслень та їх трансформації.

*Метою дисертаційної роботи є розвиток існуючих та побудова нових моделей розпізнавання графічної інформації на основі застосування нейронної мережі в задачах розпізнавання креслень та її подальшої трансформації з 2D зображень у 3D модель.*

*Об'єкт дослідження – процеси інтелектуального розпізнавання креслень та їх трансформації .*

*Предмет дослідження – методи та моделі розпізнавання креслень та їх трансформації*

*Наукова новизна отриманих результатів.* У дисертаційній роботі отримані такі наукові результати:

**вперше**

- визначено генезу інформаційної технології розпізнавання графічної інформації як комплекс методик на основі концепції нейромережевого підходу, що застосовуються для її перетворення із двовекторного простору у тривекторний з можливістю перенесення точних деталей креслення у 3D модель;
- розроблено модель класифікатора графічних зображень на основі покриття класів та елементарних класифікаторів примітивів для підвищення ефективності навчання згорткової нейронної мережі.

**удосконалено**

- архітектуру згорткової нейронної мережі, яка, на відміну від чинних рішень, враховує логічні процедури навчання на основі пошуку інформативних фрагментів описів для інтелектуального визначення наявності або відсутності конкретних графічних примітивів у системі розпізнавання креслень;
- контекстну модель інформаційної технології розпізнавання графічної інформації, яка, на відміну від інших, аналізує взаємодію процесів всередині нейронної мережі та зовнішні впливи на неї, що дозволяє представити інформаційну технологію як в агрегованому системному контексті (макрорівень), так і в деталізованому контексті у вигляді простих проєкцій (мікрорівень) із визначенням впливу рівнів декомпозиції на процес прийняття рішень щодо розпізнавання графічної інформації.

**набуло подальшого розвитку**

- систематизація та узагальнення методів, моделей, алгоритмів та програмного забезпечення задачі розпізнавання графічної інформації, що, на відміну від існуючих підходів, базуються на

методах штучного інтелекту і дозволяють використовувати згорткові нейронні мережі в задачі розпізнавання креслень;

- етапи трансформації 2D зображень у 3D модель, які є узагальненням моделі автоматизованої підготовки виробництва для задачі розпізнавання креслень, і сприяють подальшій автоматизації проектно-конструкторських та проектно-технологічних робіт;
- принципи конструювання логічних процедур навчання нейронної мережі з використанням апарату логічних функцій, які, на відміну від існуючих рішень, на етапі конструювання множини елементарних класифікаторів для креслярських примітивів зводяться до знаходження допустимих та максимальних кон'юнкцій для функції опису класу графічних об'єктів, що дозволить на практиці створювати ефективні програмні рішення CNN у задачах розпізнавання креслень та їх трансформації.

*Практична значимість наукових результатів.* Спроектвана в роботі засобами CASE-технології ERwin інформаційна технологія розпізнавання графічної інформації на основі нейронної мережі, розпізнавання креслень та трансформації з визначеними зовнішніми впливами на неї та взаємодії процесів всередині системи дає можливість програмно реалізувати програмні застосунки для розв'язання завдань розпізнавання графічної інформації на основі нейронної мережі та перетворення інформаційних масивів із двовекторного простору у тривекторний. Розроблені алгоритми та архітектура програмного додатку, які базуються на створеній моделі класифікатора графічних зображень на основі покриттів класів та ЕК примітивів для підвищення ефективності навчання CNN, дозволяють в 1,5-2 рази зменшити обчислювальні витрати на навчання CNN та до 2 разів зменшити сумарну похибку навчання CNN, що забезпечує зменшення ресурсоемності та похибки розпізнавання графічної інформації на основі

нейронної мережі та перетворення інформаційних масивів із двовекторного простору у тривекторний.

Результати інформаційної технології розпізнавання графічної інформації на основі нейронної мережі апробовані і впроваджені в наступних продуктових компаніях: ТОВ «ФБЮЧЕ КОНСАЛТИНГ», ТОВ «ІНТЕРНЕТ ІНВЕСТИЦІЙНА ГРУПА».

Дисертаційна робота виконана відповідно до плану науково-дослідних робіт Державного торговельно-економічного університету. НДР № 0122U001549 «Моделювання інтелектуальних систем управління діяльністю підприємств», (довідка від 29.08.2023 №1453/24), виконавцем якої є здобувач, а її результати включають наукові дослідження зазначеної теми.

Результати дисертаційного дослідження використано у навчальному процесі Державного торговельно-економічного університету.

*Ключові слова: інформаційна технологія, графічна інформація, зображення, нейронна мережа, розпізнавання, модель, зображення, образ.*

## SUMMARY

Khorolska K.V. Information technology for graphic information recognition based on neural networks. - Qualification work in manuscript form.

Thesis for the degree of Doctor of Philosophy in the field of "Information Technology" by specialty 122 "Computer Science". State University of Trade and Economics, Kyiv, 2023.

The thesis is a comprehensive study on the modeling, development, and application of information technology for graphic information recognition based on neural networks.

The relevance of the research topic is determined by the importance of recognizing graphical information and the practical application of the theoretical foundations for recognizing graphical information based on highly effective solutions. At this stage of practical application development, several unresolved issues remain in the theoretical foundations of graphical information recognition: quality contour extraction; automated object extraction; variability in spatial placement and recognition of objects; practical application of methodological developments in recognizing graphical information; and high-quality classification of recognition objects.

Visual image recognition is a crucial component of management and information processing systems, automated systems, and decision-making systems. Tasks associated with classifying and identifying objects, phenomena, and signals, characterized by a finite set of certain properties and features, arise in fields such as robotics, information retrieval, monitoring and analysis of visual data, and artificial intelligence research. Algorithmic processing and classification of images are applied in security systems, access control and management, virtual reality systems, and information search systems. Moreover, with the broad implementation of virtual reality systems and the development of the gaming industry, and considering that a 3D model is initially created as its 2D version, an acute need arises for the rapid conversion of two-dimensional images into three-dimensional models.



A significant role in solving the mentioned problems is played by the creation of software complexes and mathematical tools based on the application of neural networks, expert systems, and cybernetics.

Particular attention is given to artificial neural networks (or simply neural networks) - computing systems that learn from observational data through an optimization process, wherein model parameters are iteratively adjusted to minimize the difference between the predicted outcome and the actual result. Neural networks, especially convolutional neural networks (CNN), are widely recognized for their effectiveness in analyzing visual images. CNNs are a class of deep learning models specifically designed for processing 2D images. They consist of multiple layers of small neuron collections that process parts of the input image called receptive fields. The outputs from these collections are arranged so they overlap for a better representation of the output image, a key feature of CNNs. Moreover, they are translation-invariant, meaning they can identify an object as the same when it appears in different representations. These functions allow CNNs to capture intricate patterns in spatial and temporal domains – a critical aspect in the task of 3D reconstruction from 2D images. Artificial neural networks also excel in processing noisy, incomplete, or ambiguous data – a scenario commonly encountered in image processing tasks. They can extract significant features even from flawed data (noise, missing values, duplicates, contradictions), ensuring the model's reliability. Additionally, neural networks, specifically CNNs, can recognize hierarchical patterns in the data. For instance, in image processing tasks, they might identify edges and color gradients at a lower level and shapes or parts of an object at a higher level. This functionality is paramount in tasks like 3D reconstruction, which require the model to detect high-level features and relationships in 2D images.

The primary focus of the thesis research is on the recognition of graphical information based on a neural network, specifically the processing of graphical images in the form of drawings and their subsequent transformation into a 3D

model using the logical learning procedures (LLP) of a convolutional neural network (CNN).

The informational base of the research consists of regulatory materials, scientific works of contemporary domestic and foreign scientists and practitioners, as well as empirical results of the author's own investigations.

The theoretical foundation of the work is a model of a graphic image classifier based on class coverings and elementary primitive classifiers to enhance the training efficiency of CNN. Additionally, the principles of constructing logical learning procedures using logical functions are explored. This will practically enable the creation of efficient CNN software solutions in tasks of recognizing drawings and their transformation.

*The goal of the doctoral thesis* work is to develop existing and construct new models for recognizing graphical information based on the application of a neural network in tasks related to recognizing drawings and their subsequent transformation from 2D images into a 3D models.

*The research object* is the processes of intelligent recognition of drawings and their transformation.

*The subject of the study* is methods and models for recognizing drawings and their transformation.

*The scientific novelty of the obtained results.* Given thesis presents the following scientific results:

*For the first time:*

- Defined the genesis of information technology for recognizing graphical information as a complex of methodologies based on the concept of the neural network approach, used to transform it from a bi-vector space to a tri-vector one with the ability to transfer precise drawing details into a 3D model.

- Developed a model classifier of graphical images based on the coverage of classes and elementary classifiers of primitives to increase the efficiency of training convolutional neural networks (CNN).

*Improved:*

- The architecture of the convolutional neural network which, unlike current solutions, considers logical training procedures based on searching informative description fragments for intelligent identification of the presence or absence of specific graphic primitives in the drawing recognition system.

- The contextual model of information technology for recognizing graphical information which, unlike others, analyzes the interaction of processes within the neural network and external influences on it. This allows presenting the information technology both in an aggregated system context (macro level) and in a detailed context in the form of simple projections (micro level) determining the influence of decomposition levels on the decision-making process regarding the recognition of graphic information.

*Further development has been achieved in:*

- Systematization and generalization of methods, models, algorithms, and software of the graphical information recognition task, which, unlike existing approaches, are based on artificial intelligence methods and allow the use of CNN in drawing recognition tasks.

- Stages of transforming 2D images into a 3D model, which generalizes the model of automated production preparation for the drawing recognition task, and further automate design and technological work.

- Principles of constructing logical training procedures of the neural network using logical functions. Unlike existing solutions, at the stage of constructing a set of elementary classifiers for drawing primitives, they boil down to finding permissible and maximum conjunctions for the class description function of graphic objects. This will allow in practice to create effective software solutions for CNN in tasks of recognizing drawings and their transformation.

*Practical significance of the scientific results.* The information technology for recognizing graphical information based on neural networks, drawing recognition, and transformation, designed in the study using CASE technology ERwin, considering external influences on it and the interaction of processes

within the system, provides the opportunity to implement software applications to solve tasks of graphical information recognition based on neural networks and converting data from bi-vector space to tri-vector space.

The developed algorithms and software application architecture, grounded on the created model of graphical image classifiers based on class coverages and elementary classifiers (EC) of primitives to enhance CNN training efficiency, allow a 1.5-2 times reduction in computational costs for CNN training and up to a 2-fold decrease in the overall training error of CNN. This ensures a reduction in resource intensity and error in recognizing graphical information based on neural networks and transforming data from a bi-vector space to a tri-vector one.

Results of the information technology for recognizing graphical information based on neural networks have been tested and implemented in the following product companies: LLC "FUTURE CONSULTING", LLC "INTERNET INVESTMENT GROUP".

The doctoral thesis work was carried out according to the research work plan of the State Trade and Economic University. R&D No. 0122U001549 "Modeling of Intellectual Management Systems in Enterprise Activities", (reference from 29.08.2023 No.1453/24). The executor of this work is the candidate, and its results include scientific research on the specified topic.

The results of the doctoral thesis research have been used in the educational process of the State Trade and Economic University.

*Keywords: information technology, graphical information, picture, neural network, recognition, model, picture, image.*

## СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА

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

1.Khorolska K., Artificial intelligence face recognition for authentication./ Kryvoruchko, O., Bebeshko, B., Khorolska, K., Desiatko, A., Kotenko, N. (2020). Technical Sciences and Technologies, 2 (20), 139-148.

*(Особистий внесок: метод отримання даних для навчання систем безпеки з соціальних мереж і взаємодії з людиною для майбутнього використання в процесі аутентифікації в різних інформаційних системах).*

2.Khorolska, K. (2022). Потенціал застосування різних методів штучного інтелекту у задачі розпізнавання креслень та трансформації 2D→3D. *Електронне фахове наукове видання "Кібербезпека: освіта, наука, техніка";*, 1(17), 21-30. <https://doi.org/10.28925/2663-4023.2022.17.2130>

3.Khorolska, K. (2022). Аналіз основних методів розпізнавання креслень та можливостей трансформації з 2D У 3D. *Електронне фахове наукове видання "Кібербезпека: освіта, наука, техніка"* 4(16), 185-193. <https://doi.org/10.28925/2663-4023.2022.16.185193>

4.Хорольська К. Аналіз основних підходів до вирішення задачі конвертації двовимірних зображень в тривимірну модель Вісник КрНУ імені Михайла Остроградського. Випуск 3/2022 (134) DOI <https://doi.org/10.32782/1995-0519.2022.3.7>

5.Bebeshko, B., Khorolska, K., Kotenko, N., Desiatko, A., Sauanova, K., Sagyndykova, S., Tyshchenko, D. 3D modelling by means of artificial intelligence (2021) Journal of Theoretical and Applied Information Technology, 99 (6), pp. 1296-1308.

*(Особистий внесок: аналіз особливостей розпізнавання 2D зображень і створення 3D моделей з використанням згортових нейронних мереж та визначенно можливі обмеження використання CNN для генерації 3D моделей з урахуванням вихідної роздільної здатності та швидкості генерації).*

6.Khorolska K., Lazorenko V., Bebeshko B., Desiatko A., Kharchenko O., Yaremych V. (2022) Usage of Clustering in Decision Support System. In: Raj J.S., Palanisamy R., Perikos I., Shi Y. (eds) Intelligent Sustainable Systems. Lecture Notes in Networks and Systems, vol 213. Springer, Singapore. [https://doi.org/10.1007/978-981-16-2422-3\\_49](https://doi.org/10.1007/978-981-16-2422-3_49)

*(Особистий внесок: побулова алгоритму на основі методу К-середніх із залежними рухами, що дозволяє конструктивно знаходити оптимальні стратегії в проектах)*

7.Lakhno, V., Akhmetov, B., Smirnov, O., Chubaievskyi, V., Khorolska, K., Bebeshko, B. (2023). Selection of a Rational Composition of Information Protection Means Using a Genetic Algorithm. In: Rajakumar, G., Du, KL., Vuppalapati, C., Beligiannis, G.N. (eds) Intelligent Communication Technologies and Virtual Mobile Networks. Lecture Notes on Data Engineering and Communications Technologies, vol 131. Springer, Singapore. [https://doi.org/10.1007/978-981-19-1844-5\\_2](https://doi.org/10.1007/978-981-19-1844-5_2)

*(Особистий внесок: побудова модуля системи підтримки прийняття рішень для вирішення задачі вибору та оптимізації кількості систем захисту інформації).*

8.Khorolska, K., Bebeshko, B., Desiatko, A., & Lazorenko, V. (2021). 3D models classification with use of convolution neural network. Paper presented at the CEUR Workshop Proceedings, 3179 25-34. [http://ceur-ws.org/Vol-3179/Paper\\_3.pdf](http://ceur-ws.org/Vol-3179/Paper_3.pdf)

*(Особистий внесок: описано мультипрезентаційну структуру класифікації 3D-моделі).*

9.Khorolska K. , Skladannyi P., Sokolov V., Korshun N., Bebeshko B., Lakhno V., Zhumadilova M (2022) Application of a convolutional neural network with a module of elementary graphic primitive classifiers in the problems of recognition of drawing documentation and transformation of 2D to 3D models. Journal of Theoretical and Applied Information Technology 31st December 2022. Vol.100. No 24 <http://www.jatit.org/volumes/Vol100No24/18Vol100No24.pdf>

*(Особистий внесок: побудовано архітектуру згорткової нейронної мережі з модулем елементарних класифікаторів графічних примітивів та розроблена модель класифікатора графічного зображення на основі охоплених класів і елементарних примітивних класифікаторів.)*

*Наукові праці, які засвідчують апробацію матеріалів дисертації:*

10. Khorolska K. Use of AI in data protection/ Kryvoruchko O., Vebeshko V., Khorolska K. // Безпека ресурсів інформаційних систем : збірник тез І Міжнародної науково-практичної конференції (м. Чернігів 16-17 квітня 2020 р.). – Чернігів : НУЧП, 2020. – с.15-18

*(Особистий внесок: описано використання технологій штучного інтелекту в основних сферах захисту даних).*

11. Бебешко Б.Т., Лазоренко В.В., Хорольська К.В. Безпека інтелектуальної системи управління цифровими активами за допомогою методу k-means при дослідженні видобутку даних // Кібергігієна. Кібербезпека. Безпека держави: матеріали наукових семінарів (Київ, 27 листопада 2020 р.)/відп. ред. АМ Десятко.–Київ: Київ. нац. торг.-екон. ун-т, 2020.–с.34-36

*(Особистий внесок: важливість використання методів нейронних мереж для виявлення закономірностей та надання їм оцінки).*

12. Лазоренко В.В., Бебешко Б.Т., Хорольська К.В. Аналіз методів прогнозування кібератак // Комплексне забезпечення якості технологічних процесів та систем (КЗЯТПС – 2021) : матеріали тез доповідей XI Міжнародної науково-практичної конференції (м. Чернігів, 26–27 травня 2021 р.) : у 2 т. / Національний університет «Чернігівська політехніка» [та ін.] ; відп. за вип.: Єрошенко Андрій Михайлович [та ін.]. – Чернігів : НУ «Чернігівська політехніка», 2021. – Т. 2. – 236 с. ISBN 978-617-7932-16-0

*(Особистий внесок: описано сьюсіб прогнозування кібератак з неповними даними за допомогою передбачення сигналу для оцінки відповідності нетрадиційних сигналів для прогнозування).*

13. Khorolska K. 3D Model reconstruction using convolutional neural networks for 2D image processing. Proceedings of the VI International Scientific and Practical Conference. Osaka, Japan. 2023. Pp.457-459 DOI: 10.46299/ISG.2023.1.6

*Наукові праці, які додатково відображають наукові результати дисертації*

14. Khorolska Karyna VR-technology as a modern architecture tool. Management of Development of Complex Systems / Tsiutsiura, Svitlana, Bebeshko, Bohdan, & Khorolska, Karyna, (2020). Management of Development of Complex Systems, 42, 69 – 74, dx.doi.org\10.32347/2412-9933.2020.42.69-74.

*(Особистий внесок: проаналізовано вплив VR на кінцевий дизайн простороої візуалізації).*

15. Хорольська К. В. UX-дизайн інформаційної системи підприємства торгівлі. / Котенко Н.О., Жирова Т.О., Десятко А.М., Хорольська К.В., Бебешко Б.Т., Тогжанова К.О. // Вісник Кременчуцького національного університету імені Михайла Остроградського. 2020. Вип. № 3 (122). С. 107–112. DOI: 10.30929/1995-0519.2020.3.67-74

*(Особистий внесок: проаналізовано інструменти для проектування клікабельного прототипу з точки зору дизайну інтерфейсу).*

16. Tetiana Zhyrova, Nataliia Kotenko, Volodymyr Tokar, Karyna Khorolska, Bohdan Bebeshko, (2021) Testing the Accessibility of Web-applications The International Scientific Journal «Computer Systems and Information Technologies» 2021, #3 DOI: <https://doi.org/10.31891/CSIT-2021-5-12>

*(Особистий внесок: здійснено ґрунтовний аналіз стандартів з питань доступності веб-додатків).*

17. Lakhno V., Akhmetov B., Ydyryshbayeva M., Bebeshko B., Desiatko A., Khorolska K. (2021) Models for Forming Knowledge Databases for Decision Support Systems for Recognizing Cyberattacks. In: Vasant P., Zelinka I., Weber



GW. (eds) Intelligent Computing and Optimization. ICO 2020. Advances in Intelligent Systems and Computing, vol 1324. Springer, Cham. [https://doi.org/10.1007/978-3-030-68154-8\\_42](https://doi.org/10.1007/978-3-030-68154-8_42)

*(Особистий внесок: описано ймовірнісні моделі для виявлення мережеских вторгнень на основі використання динамічних байєсівських мереж).*

18. Lakhno, V., Mazaraki, A., Kasatkin, D., Kryvoruchko, O., Khorolska, K., Chubaievskiy, V. (2023). Models and Algorithms for Optimization of the Backup Equipment for the Intelligent Automated Control System Smart City. In: Ranganathan, G., Fernando, X., Rocha, Á. (eds) Inventive Communication and Computational Technologies. Lecture Notes in Networks and Systems, vol 383. Springer, Singapore. [https://doi.org/10.1007/978-981-19-4960-9\\_57](https://doi.org/10.1007/978-981-19-4960-9_57)

*(Особистий внесок: проаналізовано алгоритми оптимізації резервного обладнання інтелектуальної автоматизованої системи управління).*

19. Zhyrova, T., Kotenko, N., Bebeshko, B., Khorolska, K., Shevchenko, S. (2022) Benchmarking between the DQL Index and the Web Application Accessibility Index using Automatic Test Tools CEUR Workshop Proceedings, 2022, 3288, pp. 110–116

*(Особистий внесок: описано інструменти автоматизованого тестування сайтів).*

20. Khorolska K. Cyberattacks prediction with incomplete data/ Bebeshko B., Khorolska K. // Безпека соціально-економічних процесів в кіберпросторі: зб. матеріалів Всеукр. наук.-практ. конф. (Київ, 27 бер. 2019 р.). – Київ : Київ. нац. торг.-екон. ун-т, 2019. – с.123-125

*(Особистий внесок: описано інноваційні методи, такі як прогнозна імпутація сигналу (PSI), щоб замінити відсутні значення хоста, підвищуючи точність прогнозування).*

21. Khorolska K. Usage of neural networks in image recognition / O. Kryvoruchko, K. Khorolska, V. Chubaievskiy. // Зовнішня торгівля: економіка, фінанси, право. – 2019. – №3 (104). – С. 83–85.

*(Особистий внесок: класифікація частин креслення)*

22. Bebeshko, B., Khorolska, K., Kotenko, N., Kharchenko, O., & Zhyrova, T. (2021). Use of neural networks for predicting cyberattacks. Paper presented at the CEUR Workshop Proceedings, , 2923 213-223. <http://ceur-ws.org/Vol-2923/paper23.pdf>

*(Особистий внесок: описано інструмент прогнозування кібератак).*

## ЗМІСТ

ВСТУП .....	20
СПИСОК УМОВНИХ СКОРОЧЕНЬ ТА ПОЗНАЧЕНЬ.....	27
РОЗДІЛ 1 ПРОБЛЕМАТИКА РОЗПІЗНАВАННЯ ГРАФІЧНОЇ ІНФОРМАЦІЇ ТА ТРАНСФОРМАЦІЇ МОДЕЛЕЙ 2D В 3D .....	28
1.1 Аналіз підходів до розпізнавання графічної інформації для трансформації моделей 2D у 3D.....	29
1.2 Потенціал застосування методів штучного інтелекту в задачі розпізнавання креслень та трансформації <b>2D</b> → <b>3D</b> .....	43
ВИСНОВКИ ДО РОЗДІЛУ 1 .....	60
РОЗДІЛ 2 ЗАСТОСУВАННЯ ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ З МОДУЛЕМ ЕЛЕМЕНТАРНИХ КЛАСИФІКАТОРІВ ГРАФІЧНИХ ПРИМІТИВІВ У ЗАДАЧАХ РОЗПІЗНАННЯ КРЕСЛЕНЬ І ТРАНСФОРМАЦІЇ МОДЕЛЕЙ 2D В 3D .....	62
2.1 Архітектура згорткової нейронної мережі з модулем елементарних класифікаторів графічних примітивів для вирішення задачі розпізнавання креслень та трансформації <b>2D</b> → <b>3D</b> .....	63
2.2. Алгоритм трансформації <b>2D</b> → <b>3D</b> на основі інформації з розпізнаного креслення.....	89
ВИСНОВКИ ДО РОЗДІЛУ 2.....	94
РОЗДІЛ 3. МОДЕЛЮВАННЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ РОЗПІЗНАВАННЯ ГРАФІЧНОЇ ІНФОРМАЦІЇ ДЛЯ СТВОРЕННЯ 3D МОДЕЛІ НА ОСНОВІ 2D ЗОБРАЖЕННЯ .....	96
3.1. Концептуальна модель інформаційної технології розпізнавання графічної інформації.....	96
3.2. Контекстна модель інформаційної технології розпізнавання графічної інформації на основі нейронної мережі.....	114
3.3. Архітектура програмного комплексу інформаційної технології розпізнавання графічної інформації на основі нейронної мережі.....	130
ВИСНОВКИ ДО РОЗДІЛУ 3 .....	140
ВИСНОВКИ .....	142
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	146
ДОДАТКИ.....	156
СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА .....	166

## ВСТУП

**Актуальність теми.** Зараз у світі помітна тенденція до відмови від паперових креслень на користь електронних, оскільки їх зручніше тиражувати, зберігати, передавати та змінювати. Але досі на підприємствах також існують масштабні архіви, для оцифрування яких, використовуючи існуючі технології будуть потрібні тисячі людино-годин. Тому автоматизація перенесення растрових креслень в електронний вигляд нині є актуальним і потрібним завданням. Серед існуючих індустріальних систем [1-10] вищеописана проблема вирішується у недостатньому обсязі. Водночас такі системи мають ряд недоліків, наприклад, перетворення растрового зображення на набір векторних примитивів є складним для редагування. Також потребують подальшого розвитку наукові дослідження, які можуть бути зосереджені на розв'язанні проблеми перетворення вже оцифрованих креслень у 3D-моделі або структури, які зрозумілі виробничому устаткуванню.

Одним з основних напрямків створення інтелектуальних систем розпізнавання графічної інформації та її подальшої трансформації, є розробка алгоритмів на основі нейронних мереж (НМ).

Застосування нейронних мереж викликає інтерес і є перспективним через те, що з їх допомогою можна реалізувати способи опрацювання інформації, аналогічні процесам функціонування людського мозку. Зокрема, при вирішенні завдань розпізнавання графічної інформації та її трансформації.

Теоретико-методологічні основи розв'язання задач із розпізнавання графічної інформації та її перетворення у тривимірну модель виділені в працях вітчизняних та закордонних науковців Назаркевич М.А., Котлика С., Соколова О. та Корнієнка Ю., Пузя Д. та Сомова С., Лахна В.А., Криворучко О.В., Роскладки А.А., Гамалія В.Ф., Ванги З., Цумури К. та Сятоа Ю.

Таким чином, посилення вимог до ефективності систем розпізнавання графічної інформації та для подальшої трансформації моделей 2D у 3D, перспективність використання нейромережових засобів для цих завдань, малодоступність практичного аспекту розпізнавання графічної інформації для критичного аналізу внаслідок широкого використання розробок рівня «ноу-хау», які захищені власниками відповідних програмних застосунків, недостатня взаємопов'язаність існуючих нейромережових методів та засобів розпізнавання графічної інформації для подальшої трансформації моделей 2D у 3D обумовлюють актуальність обраної теми дисертаційного дослідження.

**Метою дисертаційної роботи** є розвиток існуючих та побудова нових моделей розпізнавання графічної інформації на основі застосування нейронної мережі в задачах розпізнавання креслень та її подальшої трансформації з 2D зображень у 3D модель.

Для досягнення поставленої мети необхідно вирішити такі завдання:

1. Провести критичний аналіз методів, моделей, алгоритмів та відповідного програмного забезпечення, що використовуються в задачі розпізнавання креслень та їх подальшої трансформації у 3D модель.

2. Оцінити потенціал використання методів штучного інтелекту в задачі розпізнавання креслень та їх трансформації у 3D модель.

3. Проаналізувати існуючі підходи до трансформації 2D зображень у 3D модель.

4. Розробити модель класифікатора графічних зображень на основі покриттів класів та елементарних класифікаторів.

5. Удосконалити принципи конструювання логічних процедур навчання з використанням апарату логічних функцій в задачах розпізнавання креслень та їх трансформації у 3D модель.

6. Розробити новий підхід до визначення генези інформаційної технології розпізнавання графічної інформації на основі використання нейронної мережі.

7. Провести декомпозицію інформаційної технології розпізнавання графічної інформації на основі нейронної мережі та визначити внутрішні й зовнішні фактори, що впливають на процес прийняття рішень в задачі розпізнавання графічної інформації.

**Об'єкт дослідження** – процеси інтелектуального розпізнавання креслень та їх трансформації .

**Предмет дослідження** – методи та моделі розпізнавання креслень та їх трансформації

**Методи дослідження.** У процесі дослідження, враховуючи особливості предметної галузі та сформульованих задач, використані методи емпіричного та теоретичного дослідження: системний підхід, моделювання (графічне, математичне (булева алгебра)), аналіз та синтез (декомпозиція, формалізація, аналогія, порівняння), semi-supervised learning (метод машинного навчання).

**Наукова новизна отриманих результатів.** У дисертаційній роботі отримані такі наукові результати:

**вперше**

- визначено генезу інформаційної технології розпізнавання графічної інформації як комплекс методик на основі концепції нейромережевого підходу, що застосовуються для її перетворення із двовекторного простору у тривекторний з можливістю перенесення точних деталей креслення у 3D модель;
- розроблено модель класифікатора графічних зображень на основі покриття класів та елементарних класифікаторів примітивів для підвищення ефективності навчання згорткової нейронної мережі.

**удосконалено**

- архітектуру згорткової нейронної мережі, яка, на відміну від чинних рішень, враховує логічні процедури навчання на основі пошуку інформативних фрагментів описів для інтелектуального

визначення наявності або відсутності конкретних графічних примітивів у системі розпізнавання креслень;

- контекстну модель інформаційної технології розпізнавання графічної інформації, яка, на відміну від інших, аналізує взаємодію процесів всередині нейронної мережі та зовнішні впливи на неї, що дозволяє представити інформаційну технологію як в агрегованому системному контексті (макрорівень), так і в деталізованому контексті у вигляді простих проєкцій (мікрорівень) із визначенням впливу рівнів декомпозиції на процес прийняття рішень щодо розпізнавання графічної інформації.

#### **набуло подальшого розвитку**

- систематизація та узагальнення методів, моделей, алгоритмів та програмного забезпечення задачі розпізнавання графічної інформації, що, на відміну від існуючих підходів, базуються на методах штучного інтелекту і дозволяють використовувати згорткові нейронні мережі в задачі розпізнавання креслень;
- етапи трансформації 2D зображень у 3D модель, які є узагальненням моделі автоматизованої підготовки виробництва для задачі розпізнавання креслень, і сприяють подальшій автоматизації проєктно-конструкторських та проєктно-технологічних робіт;
- принципи конструювання логічних процедур навчання нейронної мережі з використанням апарату логічних функцій, які, на відміну від існуючих рішень, на етапі конструювання множини елементарних класифікаторів для креслярських примітивів зводяться до знаходження допустимих та максимальних кон'юнкцій для функції опису класу графічних об'єктів, що дозволить на практиці створювати ефективні програмні рішення CNN у задачах розпізнавання креслень та їх трансформації.

**Особистий внесок здобувача.** Наукові результати дослідження, які виносяться на захист, одержані автором самостійно.

**Зв'язок роботи з науковими програмами, планами, темами.** Дисертаційна робота виконана відповідно до плану науково-дослідних робіт Державного торговельно-економічного університету. НДР № 0122U001549 «Моделювання інтелектуальних систем управління діяльністю підприємств», (довідка від 29.08.2023 №1453/24), виконавцем якої є здобувач, а її результати включають наукові дослідження зазначеної теми. Особистий внесок автора полягає в оціненні потенціалу використання методів штучного інтелекту в задачах розпізнавання технічної документації підприємства та розробці нового підходу до визначення інформаційної технології розпізнавання графічної інформації у технічній документації підприємства на основі використання нейронної мережі.

**Практична значення отриманих наукових результатів.** Спроектвана в роботі засобами CASE-технології ERwin інформаційна технологія розпізнавання графічної інформації на основі нейронної мережі, розпізнавання креслень та трансформації з визначеними зовнішніми впливами на неї та взаємодії процесів всередині системи дає можливість програмно реалізувати програмні застосунки для розв'язання завдань розпізнавання графічної інформації на основі нейронної мережі та перетворення інформаційних масивів із двовекторного простору у тривекторний. Розроблені алгоритми та архітектура програмного додатку, які базуються на створеній моделі класифікатора графічних зображень на основі покриттів класів та елементарних класифікаторів примітивів для підвищення ефективності навчання CNN, дозволяють в 1,5-2 рази зменшити обчислювальні витрати на навчання CNN та до 2 разів зменшити сумарну похибку навчання CNN, що забезпечує зменшення ресурсоемності та похибки розпізнавання графічної інформації на основі нейронної мережі та перетворення інформаційних масивів із двовекторного простору у тривекторний.



**Публікації.** Наукові результати та висновки дисертаційного дослідження підтверджуються публікаціями у наукових виданнях та їх апробацією на науково-практичних конференціях. Опубліковано 22 праці, з них: 8 статей опубліковано у фахових виданнях України категорії «Б», 1 в міжнародному фаховому виданні, 8 статей проіндексовані в міжнародній науково-метричній базі Scopus та 5 тез доповідей на всеукраїнських і міжнародних наукових конференціях.

**Апробація результатів дослідження.** Матеріали дисертаційного дослідження обговорювалися на всеукраїнських і міжнародних наукових конференціях. Зокрема: IEEE 8th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T) (Харків, 2021), 6th International Conference on Inventive Communication and Computational Technologies, ICICCT 2022 (м. Таміл Наду, Індія, 2022), Intelligent Communication Technologies and Virtual Mobile Networks (м. Таміл Наду, Індія, 2022), 2022 Workshop on Cybersecurity Providing in Information and Telecommunication Systems, CPITS 2022 (м.Київ, Україна, 2022), Intelligent Sustainable Systems (м. Тірунелвелі, Індія, 2021), 8th International Scientific Conference «Information Technology and Implementation» Workshop, IT and I-WS 2021 (м.Київ, Україна, 2021), 2021 Workshop on Cybersecurity Providing in Information and Telecommunication Systems, CPITS 2021 (м.Київ, Україна, 2021), VI International Scientific and Practical Conference (м. Осака, Японія, 2023), Всеукраїнській науково-практичній конференції «Нові інформаційні технології управління бізнесом» (м.Київ, Україна, 2021), Дев'ятій міжнародній науково-практичній конференції «Комплексне забезпечення якості технологічних процесів та систем» (м. Чернігів, Україна 2021), Першій міжнародній науково-практичній конференції «Безпека ресурсів інформаційних систем» (м. Чернігів, Україна 2021), Всеукраїнській науково-практичній конференції «Безпека соціально-економічних процесів в кіберпросторі» (м.Київ, Україна, 2019).

**Обсяг та структура роботи.** Дисертація складається зі вступу, трьох розділів, висновків, списку використаних джерел і додатків. Дисертаційна робота має 29 рисунків, 7 таблиць, 5 додатків. Список використаних джерел містить 95 найменувань. Загальний обсяг роботи складає 171 сторінка, обсяг основного тексту – 126 сторінок.

## СПИСОК УМОВНИХ СКОРОЧЕНЬ ТА ПОЗНАЧЕНЬ

- IT – інформаційні технології;
- 2D (two dimensional) – двовимірний;
- 3D (three dimensional) – тривимірний;
- ШІ – штучний інтелект (AI – artificial intelligence);
- ПЗ – програмне забезпечення;
- ПП – програмний продукт;
- 2D→3D – конвертація/трансформація 2D у 3D;
- САПР – система автоматизованого проєктування;
- CSG - constructive solid geometry - конструктивна геометрія твердого тіла;
- B-rep - bounding representation - обмежувальне представлення;
- ШНМ - штучна нейронна мережа;
- CNN - convolutional neural network – згорткова нейронна мережа;
- FCN – fully convolutional network – повністю згорткова мережа;
- CBR- case-based reasoning - міркування на основі прецедентів
- ЕО – еволюційні обчислення;
- НЛ – нечітка логіка;
- СМ – семантична мережа;
- НМ – нейронна мережа;
- ЕК – елементарний класифікатор;
- ОВН – об’єкт, що використовується для навчання;
- ЛПН – логічна процедура навчання;
- НВ – навчальна вибірка;
- ОП – описові показники;
- ОМ – опорна множина;
- АРО – алгоритм розрахунку оцінок;
- ДДНФ - досконала диз’юнктивна нормальна форма;
- КФ - кон’юнктивна функція;
- ДНФ - диз’юнктивна нормальна форма;

# РОЗДІЛ 1

## ПРОБЛЕМАТИКА РОЗПІЗНАВАННЯ ГРАФІЧНОЇ ІНФОРМАЦІЇ ТА ТРАНСФОРМАЦІЇ МОДЕЛЕЙ 2D В 3D

Із розвитком інформаційних технологій (ІТ) та їх інтеграції у виробництво (незалежно від цільової спрямованості – машинобудування, розробка ігор, легка промисловість, виготовлення меблів, медицина та ін.) можна обґрунтовано говорити про глобальний тренд щодо відмови від традиційних паперових носіїв конструкторської документації, наприклад, креслень. Перехід на електронні носії не тільки сприяє зручнішим процедурам, пов'язаним з редагуванням, тиражуванням, зберіганням, передачею оцифрованих даних, але й дозволяє оперативніше переходити на нові технології. Це повною мірою стосується і задач переходу від 2D моделювання графічного об'єкта до 3D моделей.

Двовимірні креслення (2D) найчастіше складні для розуміння. Конструкторам не завжди зручно вносити корективи до таких 2D моделей. Відповідно, створення систем на основі інформаційних технологій, для автоматичної реконструкції (розпізнавання) або трансформації 2D моделей в 3D моделі або  $2D \rightarrow 3D$ , дозволить багаторазово скоротити не тільки трудомісткість і час проектування нових графічних об'єктів або модернізацію існуючих, але дозволить візуалізувати зовнішній вигляд графічного об'єкта потенційного замовника або покупця, що не має глибоких знань у сфері просторової інтерпретації 2D моделей.

Як зазначає ряд дослідників: Котлик С., Соколова О. та Корнієнко Ю., Назаркевич М.А., Пузь Д., Кіктєв М.О. та Сомов С. [1, 2], і враховуючи той факт, що підприємства та організації накопичили величезні масиви креслярсько-конструкторської документації, розвиток та вдосконалення методів, моделей та ІТ, що сприяють автоматизації розпізнавання та подальшої трансформації/конвертації 2D у 3D моделі є актуальним науково-технічним завданням. Не варто також забувати, що сама по собі процедура

оцифрування наявної традиційної паперової документації є не тільки трудомістким процесом, що займає сотні, а часом і тисячі людино-годин, але він також пов'язаний з необхідністю вирішення і паралельних завдань. До подібних завдань належить, наприклад, необхідність захисту отриманих 3D моделей, якщо йдеться про моделі графічних об'єктів, що стосуються оборонної сфери, або вони є інтелектуальною власністю власника.

Слід зазначити, що без сучасних технологій та систем штучного інтелекту (ШІ/AI – artificial intelligence) вирішити подібне завдання складно, що саме собою породжує нове завдання даного дисертаційного дослідження - розвиток математичних моделей для штучного інтелекту в задачах розпізнавання 2D зображень для їх подальшої конвертації в 3D модель.

### **1.1 Аналіз підходів до розпізнавання графічної інформації для трансформації моделей 2D у 3D.**

Ідея автоматизації процедур трансформації  $2D \rightarrow 3D$  не є новою. Власне, ще на початку 60-х років минулого сторіччя з розвитком програмного забезпечення (ПЗ) для машинобудування стали з'являтися наукові роботи, присвячені даній проблематиці. Піонером у цьому питанні вважаються розробники фірми «SketchPad», (1963 р. США) [3, 4]. Саме вони запропонували на ринку перший програмний продукт (ПП) для інтерактивного проєктування графічних об'єктів, які використовуються у літальних апаратах. По суті, саме розробники «SketchPad» стали одними з перших розробників систем автоматизованого проєктування – САПР. Сьогодні аналогічних програмних продуктів налічується кілька сотень. Огляд та аналіз переваг та недоліків деяких з них буде наведено в цьому розділі.

У період з 60-х по 90-і розробників і споживачів очевидно влаштовували можливості «SketchPad», тому особливих інновацій у цей період не спостерігалось. Зокрема, це можна пояснити і тим фактором, що

на той момент наявні можливості для конвертації  $2D \rightarrow 3D$  були достатніми, а технології штучного інтелекту в подібних завданнях ще не були затребувані. Черговий сплеск у розвитку САПР бере початок у середині 90-х років минулого сторіччя. Саме з середини 90-х років на ринку стали з'являтися системи, власне, 3D моделювання та конвертування  $2D \rightarrow 3D$  [5-7].

Перші наукові основи конвертування  $2D \rightarrow 3D$  закладалися ще у 70-80-х роках минулого століття [8, 9]. Як було показано в даних дослідженнях, візуалізацію 3D об'єктів можна починати за умови наявності геометричної моделі. Такою моделлю, власне, і є креслення відповідного графічного об'єкта. Однак при цьому виникають такі технологічні проблеми:

- 1) введення даних щодо геометричних параметрів об'єкта;
- 2) відокремлення даних від «шуму»;
- 3) формування за даними відповідної 3D моделі.

Якщо узагальнити наукові та технологічні підходи у контексті розв'язання задачі трансформації  $2D \rightarrow 3D$ , то можна виділити дві групи [10]:

- 1) Підхід CSG (constructive solid geometry);
- 2) Підхід B-rep (bounding representation).

При першому підході в основу трансформації  $2D \rightarrow 3D$  ставиться припущення, що тривимірні об'єкти можуть будуватися з двовимірних примітивів на основі ієрархічного способу [11]. У цьому безпосередньо креслення 2D моделей є як певний набір даних (DataSet), що виступають базою даних (БД). Таким чином, ця БД містить шаблони для перекладу моделі у формат 3D. У процесі трансформації або конвертації  $2D \rightarrow 3D$  примітиви (прямі, кола, прямокутники, трапеції та ін.) збирають у 3D модель. Математичною основою такого синтезу виступають булеві операції. Однак цей підхід не позбавлений недоліків. Головним з них є те, що підхід CSG не дає задовільних результатів при розпізнаванні складних креслень.

Крім того, якщо поверхня на 3D моделі буде складною, то базові примітиви не дозволяють відобразити всі особливості графічного об'єкта.

Розглянемо підхід V-рег. Тут є власні алгоритми, які забезпечують поетапну генерацію тривимірних вершин з урахуванням креслення з наступним синтезом ребер з урахуванням координат раніше згенерованих вершин. При V-рег підході забезпечується вища точність тривимірного моделювання об'єкту. Проте V-рег підхід спочатку передбачає використання великих обчислювальних ресурсів, і навіть великих обсягів машинної пам'яті. Крім того, як зазначають Брайтенберг М., Жу Х. та Мен Ч. в роботах [12, 13] 3D модель, створена при V-рег підході, може давати неоднозначні конструкції та хибні елементи.

У 1973 р. Ідесава М. [14, 15] запропонував алгоритм та низку критеріїв для видалення хибних елементів при трансформації  $2D \rightarrow 3D$ . Так, наприклад, як помилкові трактуються тривимірні вершини, які належать до єдиної поверхні.

У 1976 р. Лафує Дж. у роботі [16] наводить опис евристичної процедури, яка дозволяла видаляти хибні елементи з 3D моделі та знаходити справжні елементи. Однак і цей підхід не позбавлений недоліків. Найсуттєвішим з них можна вважати неможливість видалення всіх хибних елементів із 3D моделі. А другим, але не менш значним недоліком вважається необхідність чіткого дотримання формату вхідних даних. Оскільки ці дані описують лінії об'єкта, або його координати, будь-яка помилка введення автоматично тягне за собою наступні помилки в процесі трансформації  $2D \rightarrow 3D$ .

На початку 80-х років минулого століття Прайс К. у своїх роботах [17, 18] запропонував не прив'язувати вхідні дані, що використовуються в ході трансформації  $2D \rightarrow 3D$  до заздалегідь заданого формату. У цих роботах [17, 18] описується алгоритм, відповідно до якого при трансформації  $2D \rightarrow 3D$  з ортогональних проєкцій об'єкта формуються видимі поверхні. Ці поверхні складаються із багатокутників. Особливістю запропонованого

автором алгоритму стало те, що 2D вершини сортувалися зі збільшенням їх координат у тривимірному просторі, тобто, відповідно за координатами  $X, Y, Z$ . Таке сортування забезпечувало скорочення часу пошуку координат ребер під час складання 3D моделі.

Приблизно в цей же період (1984 р.) інші автори Харалік Р. та Квіні Д. [19] запропонували свій алгоритм трансформації  $2D \rightarrow 3D$ . Алгоритм дозволяє формувати поверхні тривимірного об'єкту на основі множини 3D точок, що виступають як вершини об'єкта. Паралельно при цьому формувалися і множини поверхонь, з яких, власне і складався об'єкт. Однак, як зазначають самі автори, алгоритм не пристосований для трансформації  $2D \rightarrow 3D$  складних об'єктів.

У той же час вийшли кілька робіт співробітників Гарвардського університету Марковський Дж. і Веслі М. [20-22], які підійшли до вирішення задачі трансформації  $2D \rightarrow 3D$  з погляду математичної формалізації цього завдання. В результаті з'явилися роботи [21-22], в яких трансформація  $2D \rightarrow 3D$  розглядається як математична операція з такими сутностями – підпростір та нескінченний простір, а також їх комбінації. Реалізація описаних авторами алгоритмів вимагала досить великих обчислювальних ресурсів і була цікава швидше, як математичне дослідження, ніж як прикладна робота, яку можна використовувати на практиці.

Розвитком робіт [20-22] стала публікація статті Сакураї Х. та Гроссарда Д. [23]. У своєму дослідженні автори скоротили кількість сутностей, що використовуються під час трансформації  $2D \rightarrow 3D$ . Їхній алгоритм містив лише звичні сутності, наприклад, плоскі проєкції циліндрів, сфер, конусів і торів. У запропонованому авторами алгоритмі були відсутні обмеження на порядок розміщення ліній, що дозволило зробити вихідний алгоритм Марковського Дж. і Веслі М. більш пристосованим до реальних умов.



В міру зростання складності вихідних моделей для трансформації  $2D \rightarrow 3D$  використання традиційних формалізованих математичних підходів вже не забезпечувало прийнятної якості і швидкості трансформації. Ці складності призвели до того, що починаючи з середини 90-х років минулого століття, низка дослідників у своїх роботах стали пропонувати методи та моделі трансформації  $2D \rightarrow 3D$ , засновані на імплементації в даний процес систем штучного інтелекту.

Так, наприклад, [24] Ванга З., Цумура К. та Сятоа Ю. представили результати своїх досліджень із застосування штучної нейронної мережі (ШНМ) у ході аналізу креслень на паперових носіях. Розроблена під час цих досліджень комп'ютерна система була здатна виконувати векторизацію вихідних даних. Автори, однак, обмежилися лише такими сутностями як лінії контурів об'єктів на кресленнях та перетину ліній. Крім того, описувана система не орієнтувалася на реальне виробництво, а швидше додавала рішення суто математичного завдання.

У подібному ключі виконані дослідження авторів публікації [25]. Вони продовжили дослідження щодо виявлення перспективності застосування штучної нейронної мережі у завданнях розпізнавання графічної інформації (креслярської документації). Однак порівняно з попередньою роботою [24], у [25] описані результати застосування ШНМ для розпізнавання символів на кресленнях. Практичні дослідження авторів були виконані лише стосовно креслень труб, що істотно знижує релевантність даного дослідження, наприклад, коли мова заходить про складні об'єкти з великою кількістю граней і поверхонь.

В [24, 26] автори описали приклад застосування багаторівневої ШНМ для трансформації  $2D \rightarrow 3D$ . Згідно з описом [24] перша нейронна мережа призначалася для отримання прямих і похилих ліній з об'єкта на кресленні, а друга нейронна мережа служила для розпізнавання інших елементів, що є на початковому зображенні, наприклад, символів.

В роботі [27] автори пропонують алгоритм, відповідно до якого вихідне креслення об'єкта трансформації  $2D \rightarrow 3D$  замінюється графом. Кожна проєкція дає компонент зв'язаності об'єкта. Якщо об'єкта на кресленні є кілька проєкцій, кожна з них відповідає компоненту зв'язності графа. Відповідно, кожен лінійний елемент на кресленні (графі) чи то криві чи відрізки можна розподілити по компонентах зв'язності. Цей алгоритм чудово показав себе в завданнях розпізнавання на полі креслення проєкційних видів, якщо їх якість і складність проєкції були відповідно високими і не складними.

Якщо говорити про методи та моделі штучного інтелекту, що застосовуються в задачі розпізнавання креслень та трансформації  $2D \rightarrow 3D$ , то зрозуміло дослідники не обмежувалися лише застосуванням нейронних мереж. Відомі й інші напрями досліджень, наприклад, застосування методів нечіткої логіки, нечітких множин та ін.

Так в роботі Ванга З. та Латіфа М. [28] під час аналізу ортогональних проєкцій об'єктів було запропоновано задіяти методи нечіткої логіки. До переваг даного підходу слід віднести можливість отримання рішення прийнятної якості для багатьох тіл, утворених компонентами призматичної форми або сформованих в результаті обертання навколо єдиної осі.

У [29, 30] розглядається метод евристичного моделювання, а також задіяння апарату нечітких множин вирішення задач з розпізнавання креслень та трансформації  $2D \rightarrow 3D$ .

У роботі [31] авторами проаналізовано особливості процедури розпізнавання об'єкта на основі його ортогональних проєкцій та врахування специфіки застосування В-гер підходу. Автори даного дослідження наводять лише приклади синтезу тривимірного об'єкта за наслідками розпізнавання ортогональних проєкцій об'єкта з плоскими гранями. Це суттєво звужує сферу застосування, запропонованої моделі.

У роботі [32] Каргасом А., Кулі П. та Річардсом Т. викладено теоретичні засади застосування алгоритмів конструктивної твердої

геометрії у задачах трансформації  $2D \rightarrow 3D$ . Ці алгоритми (англ. CSG Constructive Solid Geometry) дозволяють виконувати перетворення даних, отриманих у результаті розпізнавання ортогональних проєкцій. Однак, недоліком цієї роботи став факт, що автори обмежуються лише об'єктами призматичної форми. Хоча методи апроксимації моделей у разі узагальнення кількох математичних підходів розширюють сферу застосування викладених алгоритмів, цього недостатньо для роботи зі складними об'єктами.

Удосконаленню існуючих та розробці нових методів, моделей та інформаційних технологій для вирішення завдання розпізнавання графічної інформації (креслярської документації) та трансформації моделей  $2D \rightarrow 3D$  присвячено достатню велику кількість наукових публікацій і в останні роки. Однак, як буде показано далі, не всі ці наукові дослідження завершуються розробкою відповідного комерційного програмного продукту. Це пов'язано як з нетривіальністю самої задачі, з точки зору математичної формалізації, так і особливостями її програмування алгоритмічними мовами. До програмних продуктів, в яких вдалося досить вдало подолати складності, що виникають, можна віднести програмні продукти фірми MATRA DATAVISION ПП CAS.CADE (Франція) [33], ПП GStarCAD 2011, AutoCAD (США) [34], а також ПП Scan2CAD [35], Inkscape [36] та низку інших. Функціонал основних програмних продуктів проаналізовано надалі.

Як показує аналіз попередніх досліджень та виконаний у даному підрозділі дисертації короткий огляд функціональних можливостей програмних продуктів для вирішення завдання розпізнавання креслень та трансформації  $2D \rightarrow 3D$ , процес розпізнавання проєкційних зображень недостатньо детермінований. Це пояснюється перш за все тим, що існує нескінченне різноманіття об'єктів моделювання, що відрізняються своєю специфікою, і, відповідно, наявними особливостями трансформації  $2D \rightarrow 3D$ .

Резюмуючи оглядову частину попередніх досліджень, можна констатувати, що синтез 3D з урахуванням проєкцій тобто, фактично, 2D моделей об'єкту, який містить деякий підсумовуючий аналіз множин графічних об'єктів. Причому даний аналіз може виконуватися з урахуванням різних математичних методів і моделей.

Слід зазначити, що існує суперечність між традиційною процедурою підготовки графічної інформації (креслярської документації) на паперових носіях до 80-90-х років 20-го століття та новими методами 3D моделювання, що отримали розвиток із середини 90-х років. Це дає дійсно безмежні можливості конструкторам готувати проєктно-технічну документацію, не зациклюючись на проблемі підготовки креслярсько-конструкторської документації та особливостях введення вихідних даних. Істотно полегшує цей процес застосування прикладного програмного забезпечення.

В міру розвитку технологій машинного проєктування та відповідних САПР кількість методів і моделей, які можуть бути потенційно використані в задачі розпізнавання креслень та трансформації  $2D \rightarrow 3D$  стрімко зростає. Тому нижче у таблиці 1.1 виконано узагальнення для найбільш затребуваних методів, моделей та алгоритмів. Ця таблиця розвиває раніше виконане оглядове дослідження у роботі [30].

В контексті виконуваного в цьому розділі дисертації огляду та аналізу досліджень та підходів у питаннях розпізнавання графічної інформації для трансформації  $2D \rightarrow 3D$  слід згадати спроби вирішити дане завдання, використовуючи альтернативні підходи. Наприклад, у роботах Сазерленда та Торнтонна [37, 38] описані моделі та методи, які на думку авторів надають можливість вирішувати завдання, пов'язані з отриманням 3D моделей альтернативними шляхами, при цьому не застосовуючи спочатку як джерело вихідних даних або посередника модель 2D зображення.

Так у роботі Сазерленда [37] викладено метод, відповідно до якого проєктувальник може генерувати вершини об'єкта моделювання за допомогою так званої операції «сколювання» 2-х проєкцій.

В праці Торнтон [38] описаний схожий метод. При цьому застосовується інтерактивна техніка для отримання тривимірного введення, якщо є двовимірні проєкції об'єкта.

Таблиця 1.1

Узагальнення найбільш затребуваних методів, моделей та алгоритмів, що використовуються в задачі розпізнавання креслень та трансформації  $2D \rightarrow 3D^*$

Автори, рік, джерело	Тип поверхонь	Вхідні вимоги	Переваги	Недоліки
1	2	3	4	5
Ідесава М., 1973 р., [15]	Площина	Наявність ліній та інших примітивів.	Простота	Не дозволяє розпізнавати креслення та трансформувати $2D \rightarrow 3D$ для об'єктів зі складними формами.
Лафує Дж., 1976 р. [15]	Площина	Наявність ліній та інших примітивів за умови вказівок отворів на об'єкті.	Простота. Гарна здатність до алгоритмізування.	Не дозволяє розпізнавати креслення та трансформувати $2D \rightarrow 3D$ для об'єктів зі складними формами.
Ву Т., Хаммер Дж. М., 1977 р., [39]	Площини та циліндри	Наявність ліній та інших примітивів.	Гарна здатність до алгоритмізування.	Не дозволяє трансформувати $2D \rightarrow 3D$ для об'єктів зі складними формами. Обмежений набір примітивів для $2D$ моделі.
Прайс К., 1981 р., [17, 18]	Площина	Наявність ліній та інших примітивів.	Відносно невеликі витрати часу на пошук координат ребер під час збору $3D$ моделі.	Не дозволяє трансформувати $2D \rightarrow 3D$ для об'єктів зі складними формами. Потрібні великі обчислювальні ресурси.
Харалік Р.М. та Квіні Д., 1984 г. [19]	Площина	Наявність ліній та інших примітивів.	Алгоритм дозволяє формувати поверхні $3D$ об'єкта на основі множини $3D$ точок, що виступають як вершини об'єкта.	Алгоритм не підходить для трансформації $2D \rightarrow 3D$ складних об'єктів.
Марковський Дж. та Веслі М.А., 1981 р., [20-22]	Площина	Наявність ліній та інших примітивів.	Трансформація $2D \rightarrow 3D$ розглядається як математична операція з такими сутностями – підпростір та нескінченний простір, а також їх комбінації.	Потрібні великі обчислювальні ресурси.

1	2	3	4	5
Сакураї Х, Гроссард Д., 1982 р., [23]	Площина	Наявність ліній та інших примітивів (плоскі проєкції циліндрів, сфер, конусів і торів)	Невелика кількість сутностей, що використовуються під час трансформації $2D \rightarrow 3D$ .	Не пристосований до реальних умов виробництва.
Ванга З., Цумура К. та Сятоа Ю., 1994 г. [24]	Площина	Наявність ліній та інших примітивів (плоскі проєкції циліндрів, сфер, конусів і торів)	Задіяння потенціалу ШІ (ШНМ) для трансформації $2D \rightarrow 3D$ .	Описана система не орієнтувалася на реальне виробництво, а скоріше додавала вирішення суто математичного завдання.
Еліан Е., Джеймісон Л., Алі-Гомбе А., 2020 г., [25]	Площина	Наявність ліній та інших примітивів	Задіяння потенціалу ШІ (ШНМ) для трансформації $2D \rightarrow 3D$ .	Практичні дослідження авторів були виконані лише стосовно креслень труб.
Зехтабан Л., Елажари О., Роллер Д., 2016, [27]	Площина	Граф	Вихідне креслення об'єкта трансформації $2D \rightarrow 3D$ замінюється графом. Алгоритм чудово показав себе в завданнях розпізнавання на полі креслення проєкційних видів, якщо їх якість і складність проєкції були відповідно високими і не складними.	Алгоритм не підходить для трансформації $2D \rightarrow 3D$ складних об'єктів.
Ванг З., Латіф М., 2007 г., [28]	Площина	Наявність ліній та інших примітивів	Застосування апарату нечіткої логіки, що забезпечує гнучкість алгоритму трансформації $2D \rightarrow 3D$ .	Не завжди висока якість виконаної трансформації $2D \rightarrow 3D$ .
Каргас А., Кулі П., Річардс Т.Х.Е., 1988 г. [32]	Призматичні тіла	Наявність ліній та інших примітивів	Алгоритми дозволяють виконувати перетворення даних, отриманих внаслідок розпізнавання ортогональних проєкцій.	Досліджувалися лише об'єкти призматичної форми.

\*Джерело: побудовано автором на основі джерел [15-32]

Проте, зауважимо, що роботи [37, 38] спрямовані на вирішення іншого завдання, а саме синтезу тривимірних моделей якщо відсутні дані для 2D моделі чи їх якість не задовольняє проєктувальника.

Слід зазначити, що для багатьох підприємств традиційні 2D системи, як і раніше, мають велике поширення і продовжують використовуватися на різних стадіях виробництва. 3D моделювання об'єктів завойовує все більшу кількість прихильників серед проєктувальників та виробників, 2D системи залишаються затребуваними через такі фактори (див. рис. 1.1):

- 1) широкі можливості введення як у ручному режимі (класична креслярська документація на паперових носіях) так і з застосуванням прикладного ПЗ двовимірної геометричної, графічної, пояснювальної, алфавітно-цифрової та іншої інформації;
- 2) широкі можливості для редагування наявної документації включаючи розробки, створені кілька десятиліть тому. Не вдаючись до використання ІТ для розпізнавання та/або оцифрування наявної документації;
- 3) наявність у підприємств особливо у сфері машинобудування великих масивів креслярської документації на паперових носіях. Причому, як це не парадоксально, у деяких випадках це більш захищений варіант зберігання, ніж креслярська документація, оцифрована та збережена на електронних носіях, до яких може бути отриманий несанкціонований доступ з боку зловмисників, конкурентів або ворожих держав;
- 4) при редагуванні інформації, поданої у 2D моделях або 2D системах, розробнику достатньо для внесення змін до проєкту використовувати відносно не складний аналітичний апарат. Наприклад, більшість таких змін передбачає лише необхідність знання основ геометричних побудов на площинах двовимірної системи координат. А якщо така документація виконана досить якісно, то й наступна процедура трансформації  $2D \rightarrow 3D$  не становитиме складності;

- 5) системи 2D надають користувачам багато можливостей та засобів із вилучення інформації, яка характеризує геометрію як об'єкта в цілому (наприклад, габаритні розміри) так і його проєкцій (наприклад, їх площі). А така інформація може бути потрібна при плануванні виробництва та розміщення на відповідних площах технологічного обладнання;
- 6) сучасні системи 2D з можливостями автоматичної або автоматизованої процедури трансформації  $2D \rightarrow 3D$  вже самі по собі, як правило, містять мінімальний набір даних для отримання 3D моделей. До таких даних можна, наприклад, віднести координати точок, що відповідають вершинам проєкцій сторін 3D об'єкта моделювання. Саме ці дані найчастіше затребувані існуючими системами автоматизованого проєктування.

Однак не дивлячись на перелічені вище фактори 3D моделі поступово завоюють свої позиції, володіючи великою кількістю переваг, про які йшлося на початку даного розділу роботи.

Зауважимо, що в більшості систем 3D (наприклад, програмні продукти Autodesk TinkerCAD, DesignSpark Mechanical, FreeCAD, Autodesk AutoCAD, ZBrush, Blender та ін.) застосовуються підходи, які дозволяють синтезувати каркасне або граничне уявлення об'єкта, що моделюється в просторі. У професійних системах (наприклад Autodesk 3ds Max, Autodesk Maya) використовують узагальнені моделі просторових об'єктів. Таке уявлення передбачає наявність у проєктувальників відповідних інформаційних масивів, які ап'орі повинні відповідати всім проєкціям об'єкта у трьох основних площинах. Однак ці професійні ПП (див. таблицю 1.2) для 3D не можуть обійтися без початкового синтезу каркасної моделі. І вже після того, як така модель синтезована або отримана в результаті розпізнавання за наявними кресленнями 3D моделі, можна переходити до побудови узагальненої моделі та операціями з такою моделлю. Власне, таке міркування і диктує різноманітність шляхів залучення інструментарію



штучного інтелекту для вирішення низки завдань, що виникають на етапі розпізнавання графічної інформації (креслярської документації) та трансформації моделей  $2D \rightarrow 3D$ . На підставі вище сказаного в наступному підрозділі дисертації проаналізуємо переваги та недоліки різних методів ШІ, які потенційно можуть бути застосовані у контексті вирішення завдань даного дослідження.



Рисунок 1.1. Фактори затребуваності 2D систем

Джерело: побудовано автором

Популярні програмні продукти з функціоналом для розпізнавання  
креслень та можливостей трансформації 2D → 3D\*

Пакет прикладного ПЗ	Доступний функціонал у контексті теми дослідження	Переваги	Недоліки
1	2	3	4
AutoCAD Raster Design Tool	Розпізнавання креслення із растрових зображень. Імпорт зображень із файлів у форматі PDF. Перетворення зображень у векторний формат	Багатий функціонал	Висока вартість пакету. Від 70 до 170 000 грн. Немає задовільного рівня розпізнавання тексту в графічній інформації (креслярській документації)
Scan2CAD	Відмінні можливості для перетворення растрових елементів на вектори. Хороше розпізнавання текстової інформації на кресленнях. Роботи з багатьма форматами вихідної документації (наприклад, BMP, JPE, JPEG, JPG, TIFF та ін.)	Багатий функціонал	Складний інтерфейс. Відсутність документації українською мовою. Висока вартість пакету. Від 2000 до 10 000 грн. в місяць
AutoDWG	ПЗ для конвертації сканованої графічної інформації (креслярської документації) у різні формати. Насамперед робочий формат DWG	Багатий функціонал. Дозволяє розпізнавати текстову інформацію та рукописні креслення. Імпортувати креслення формати DWG та DXF для подальшої роботи або наступної векторизації. Відносно низька вартість	Немає документації українською мовою
Inkscape	ПЗ для роботи з векторною графікою	Можливості редагування зображень у популярних форматах SVG, PDF, PNG та ін. Розповсюджується безкоштовно	Досить простий інтерфейс. Досить обмежені можливості для вирішення завдань, пов'язаних із розпізнаванням креслень та можливостями трансформації 2D → 3D. Для цього може знадобитися додаткове комерційне програмне забезпечення
Bentley Descartess	ПЗ для розширеної обробки 3D-зображень	Можливості візуалізації відображенням об'єкта в 3D. А також поєднання об'єкта з даними CAD, BIM та GIS систем. Дозволяє	Висока вартість. В Україні практично не використовується

1	2	3	4
		конвертувати растрові документи у векторні креслення. Робота з популярними форматами PDF, IMG, JPEG, TIFF та ін.	

*\*Джерело: побудовано автором на основі джерел [33-37]*

## **1.2 Потенціал застосування методів штучного інтелекту в задачі розпізнавання креслень та трансформації $2D \rightarrow 3D$**

В міру зростання кількості та складності завдань, що виникають при оцифруванні існуючої на паперових носіях креслярсько-технічної документації, та паралельної необхідності трансформації  $2D \rightarrow 3D$  для візуалізації у 3D складних об'єктів, дослідники звернули увагу на можливості застосування технологій та систем штучного інтелекту у процесах розпізнавання креслень та трансформації  $2D \rightarrow 3D$ .

Не ставлячи за мету проведення детального огляду та аналізу всіх робіт у цьому напрямі, оскільки їх кількість постійно зростає, зупинимось на знакових публікаціях, присвячених потенціалу застосування різних методів штучного інтелекту в задачі розпізнавання креслень та трансформації  $2D \rightarrow 3D$ .

Перші дослідження, присвячені застосуванню штучного інтелекту в задачах розпізнавання зображень на кресленнях, почали з'являтися ще на початку 90-х років 20-го століття.

У роботах [40, 41] були описані ранні моделі для вирішення задач розпізнавання намальованих від руки графічних символів і примітивів на інженерних кресленнях. Автори пропонували представляти атрибутивними реляційними графами графічні символи та примітиви типових зображень, наприклад, кіл, ліній та ін. І хоча результати даного дослідження загалом були обнадійливими, описаний метод скоріше був орієнтований на

розпізнавання символічної інформації, ніж працювати безпосередньо із зображеннями на кресленні.

В працях [41, 42] було показано, що класифікація, локалізація/виявлення та сегментація є найбільш поширеними методами роботи із зображеннями на рисунках та кресленнях.

У роботі [43] Кастеллано та Вессіо представили огляд з ретроспективи застосування найбільш важливих підходів глибокого навчання до вилучення та розпізнавання образів на кресленнях/рисунках. І хоча цінність цієї роботи незаперечна з погляду глибини та хронології розвитку систем ШІ в завданнях розпізнавання рисунків/креслень, вони скоріше орієнтовані на образотворче мистецтво, ніж на роботу з креслярсько-технічною документацією. Хоча загалом паралелі очевидні.

У роботах [43-45] показано, що саме з появи згорткових нейронних мереж (CNN) відбулася революція у обробці зображень. Саме CNN стали домінувати над ринком додатків, використовуваних для систем комп'ютерного зору.

І хоча згорткові нейронні мережі, безумовно, лідирують за багатьма показниками в них, як і в інших нейронних мережах, залишається істотна вада, а саме необхідність навчання такої мережі, що не завжди можливе або пов'язане з витратами ресурсів (часових, фінансових та ін.).

Це стимулює дослідників шукати альтернативні методи та моделі для систем розпізнавання зображень на кресленнях.

Наприклад, у статті [46] розглядаються можливості застосування семантичної сегментації для вирішення завдач комп'ютерного зору та розпізнавання. Авторами показано, що семантична сегментація не представляє об'єкт у статичному діапазоні, такому як обмежувальна рамка (контури об'єкта), але дозволяє розділяти цільову область пікселів на значні одиниці при виявленні примітивів, наприклад ліній на зображенні. Отже, семантична сегментація, на думку авторів, є підходящою технікою розуміння відносин між елементами креслення.

Со Джіхо, Хеджін Парк і Сеунгьон Чу в роботі [46] показали, що багато підходів до семантичної сегментації засновані на CNN (згорткової нейронної мережі) і FCN (fully convolutional network – повністю згорткова мережа).

Зрозуміло, не відкидалися й інші методи штучного інтелекту.

Наприклад, у роботах [47, 48] пропонувалося застосовувати методи та алгоритми для еволюційного моделювання у завданнях розпізнавання зображень. Автори даних робіт акцентують увагу, що представлення форми – це важлива проблема у сфері розпізнавання образів на зображеннях. Замкнений контур, який відноситься до контуру форми, зокрема об'єкта на кресленнях, відіграє вирішальну роль у порівнянні. А оскільки контури об'єктів апіорі є стійкими і, як правило, добре помітними їх доцільно включити у процес розпізнавання. Задіяння потенціалу генетичного алгоритму дозволяє виявляти найпоширеніші фрагменти контуру. Відповідно, ці фрагменти можна використовувати для подання контурів категорії форми об'єкта.

Цікавим є й ряд робіт, наприклад, [49, 50] у яких завдання розпізнавання зображень та трансформації  $2D \rightarrow 3D$  пропонується вирішувати на основі використання апарату нечіткої логіки. На думку авторів, застосування теорії нечіткої логіки дозволить досягти хороших результатів у такому завданні як реконструкція твердотільних моделей з інженерних креслень. Як показано Латіфом та Вангом у роботі [50] інженерні креслення містять двомірні (2D) орфографічні проєкції, які завжди однозначно характеризують об'єкт загалом. І, відповідно, потрібно досить багато ітерацій і комбінаторних переборів для реконструкції тривимірного (3D) відображення об'єкта. Автори запропонували алгоритм, побудований на апараті нечіткої логіки для визначення категорії об'єкта та реалізації подальших операцій. Зокрема, показано, що як тільки об'єкт класифікується відповідною програмою як обертовий або призматичний,

основні операції обертання або видавлювання буде реалізовано для синтезу 3D-моделі.

У контексті застосування систем та методів штучного інтелекту для розпізнавання графічної інформації є й роботи, які стосуються деяких аспектів застосування такого апарату штучного інтелекту, як міркування на основі прецедентів. Наприклад, в праці [51] розглядаються можливості застосування міркувань на основі прецедентів (CBR- case-based reasoning) для визначення особистих стилів окремих графічних зображень. Відповідно до запропонованого методу, спочатку створюється серія прецедентів для намальованого від руки ескізу. Результуючий рисунок синтезується поступово з ітеративним циклом пошуку та адаптації варіантів-кандидатів доти, доки не буде досягнутий бажаний естетичний стиль.

Не розвиваючи далі оглядову частину роботи, систематизуємо основні методи штучного інтелекту, які потенційно можуть використовуватись у задачі розпізнавання зображень та трансформації  $2D \rightarrow 3D$  у табличній формі, акцентуючи увагу на перевагах та недоліках цих методів, див. таблицю 1.3.

Таблиця 1.3

Порівняльний аналіз основних методів штучного інтелекту для завдання розпізнавання креслень та можливостей трансформації

$2D \rightarrow 3D^*$

Метод ШІ	Застосування описано у роботах	Переваги	Недоліки
1	2	3	4
Міркування на основі прецедентів	[51]	1.Можливе прийняття рішень для ситуацій із високим ступенем невизначеності. 2.Є можливість пояснення отриманих рішень. 3.Гарне навчання таких систем.	1.Метод добре зарекомендував себе лише в областях, котрим властивий принцип регулярності. А за наявності великої різноманітності прецедентів трудовитрати невиправдано великі. 2.Некомпактна форма подання та зберігання знань (або досвіду).

1	2	3	4
		4.Прецеденти дозволяють врахувати суб'єктивність рішень.	3.Специфічність та трудомісткість процесів, пов'язаних із пошуком пошуку аналогічних рішень, а також складності з адаптацією рішень.
ШНМ (штучна нейронна мережа або нейронна мережа)	[24, 26, 43-45]	1.ШНМ мають хорошу стійкість до шумів вхідних даних. 2. Добре адаптуються до змін чи модифікації. 3.Мають хороші характеристики по відмовостійкості. 4. Висока швидкодія. 5.Існують гарні можливості для паралельного виконання кількох операцій одночасно.	1.ШНМ дозволяють знаходити лише субоптимальні рішення. Не мають потрібної високої точності розпізнавання. 2.Порівняно високі показники щодо вартості процесу навчання. 3.Високі часові витрати навчання. 4.Необхідність у достатніх обсягах пам'яті, особливо у випадках роботи з графічною інформацією.
Еволюційні обчислення (зокрема генетичні алгоритми)	[47, 48]	1.Добре зарекомендували себе для завдань пошуку рішень з великою розмірністю. 2.Характеризуються відсутністю обмежень на тип цільової функції. 3.Можна інтегрувати методи та моделі еволюційних обчислень (ЕО) в інші методи. Або поєднувати ЕО з методами та моделями нечіткої логіки, ШНМ та ін., що дозволяє створювати гібридні системи ШІ.	1.ЕО дозволяють знаходити лише субоптимальні рішення. Не мають потрібної високої точності розпізнавання. 2.Відносно високі показники з обчислювальної трудомісткості ЕО. 3.Складнощі з самоадаптацією в оригінальних завданнях, наприклад, коли об'єкт розпізнавання не є типовим.
Теорія нечіткої логіки (НЛ)	[26-30]	1.Системи розпізнавання графічної інформації, спроектовані з урахуванням застосування НЛ досить гнучкі. 2.Можливість оперативно змінювати правила. 3.Хороші можливості трактування неточних даних з креслення. Або для ситуацій, у яких має місце помилкова інформація. 4.Добре зарекомендували себе у складних ситуаціях,	1.Відсутня стандартна методика проектування систем розпізнавання графічної інформації, збудованих на апараті НЛ. 2.У разі збільшення кількості змінних відповідно експоненційно збільшується складність обчислень. 3.Великі трудовитрати розробки правил для таких систем.

1	2	3	4
		пов'язаних із розв'язанням задач трансформації 2D → 3D.	
Семантичні мережі (СМ)	[46]	1.Універсальність СМ. 2.Наочність уявлення системи знань. 3.Близькість структури СМ природною мовою.	1.Важко модифікувати СМ. 2.СМ пасивні структури. Що потребує розробки окремого апарату формалізації виведення. 3. Трудомісткість процедури трактування та модифікації знань, описаних на основі СМ. 4.Відсутність можливостей до самонавчання.

*\*Джерело: побудовано автором на основі джерел [24, 26, 43-51]*

Проаналізувавши підходи, які використовуються при вирішенні задач синтезу об'єктів за наявними 2D проєкціями, можна зробити такі висновки:

- 1) найбільших успіхів різними авторами було досягнуто для варіанта, коли використовувався підхід «знизу-вгору»;
- 2) найбільшу ефективність вдалося отримати на перетині класичних підходів та евристичних методів.

Розглядаючи попередні дослідження та проаналізувавши публікації за останні кілька років, можна констатувати, що у класичних підходах трансформації 2D у 3D присутні такі етапи (рис.1.2.).

1.*Етап аналізу проєкцій для 2D.* На цьому етапі виконується аналіз 2D проєкцій об'єктів. Вони можуть бути отримані, наприклад, з фотографій або рисунків, креслярсько-технічної документації. Аналізуючи ці проєкції, отримуємо геометричну інформацію про об'єкти у 2D просторі.

2.*Етап створення масивів, що містять тривимірні координати (або так звані - точкові моделі об'єктів).* На даному етапі створюються масиви, які містять тривимірні координати точок об'єктів. Ці масиви можуть бути розташовані у тривимірному просторі і представляти точкові моделі об'єктів.

3.*Етап створення попередніх каркасних моделей об'єктів.* Використовуючи масиви тривимірних координат, створюються попередні



каркасні моделі об'єктів. Ці моделі представляють собою структури, що складаються з взаємопов'язаних ліній і точок, утворюючи каркас об'єкта без поверхонь.

*4.Етап аналізу хибних геометричних елементів.* На цьому етапі є додаткове підзавдання, що зводиться до необхідності уточнення каркасних моделей об'єктів. Даний етап характеризується наявним аналізом хибних геометричних елементів, які можуть бути виявлені в попередніх каркасних моделях. Додатково проводиться уточнення каркасних моделей для виправлення цих помилок і поліпшення геометричної точності.

*5.Етап створення граничних моделей.* Після уточнення каркасних моделей виконується створення поверхонь, що складаються з граней, і формується гранична модель об'єкта. Грані можуть бути плоскими полігонами або складатися з кривих поверхонь.

*6.Етап створення конструктивних моделей.* На цьому етапі розширюється гранична модель, додаванням деталей і структур для створення конструктивних моделей об'єктів. Ці моделі мають більшу деталізацію і можуть включати додаткову інформацію, наприклад, про матеріали, текстури або фізичні властивості.

*7.Етап візуалізації.* На останньому етапі виконується візуалізація конструктивних моделей об'єктів, що дозволяє їх відображати у тривимірному просторі. Це може бути досягнуто, наприклад, за допомогою комп'ютерної графіки, рендерингу або використанням спеціального обладнання.

Перехід від третього етапу до етапів 4 і 5 по суті зводиться до необхідності розв'язання задачі трансформації каркасних моделей в граничні. Таким чином, ця підзадача може бути віднесена до більш глобальної проблеми. Цю проблему можна окреслити як завдання відновлення оригіналу за наявними проєкціями.

Створена схема показує послідовність етапів, що здійснюються для перетворення 2D інформації у 3D моделі об'єктів у класичних підходах.

Кожен етап додає нові елементи інформації та доповнює попередні, що призводить до отримання тривимірних моделей з більшою деталізацією.

Також слід зазначити, що багато комп'ютерних систем геометричного моделювання, наприклад, такі як, AutoCAD, можуть працювати як з каркасними моделями, так і з граничними. До того ж, в AutoCAD, крім роботи з каркасними та граничними моделями окремо, можливе перетворення цих моделей між собою.



**Рисунок 1.2. Модель трансформації 2D зображень у 3D модель в класичних підходах**

*Джерело: побудовано автором*

У низці досліджень, зокрема [50, 51], акцентується увага на тісному взаємозв'язку завдань розпізнавання креслень, трансформації  $2D \rightarrow 3D$  та підзавдання підготовки виробництва. Цей взаємозв'язок, як показано в [50, 51], може бути описаний такою послідовністю етапів.

1. Етап, на якому визначається система координат графічного об'єкта.
2. Етап, на якому визначаються габаритні розміри графічного об'єкта за відповідними осями.
3. Етап, на якому реалізується поділ на елементарні поверхні та відповідні примітиви.
4. Етап, на якому виділяють основні елементи графічного об'єкта та визначають його межі.
5. Етап, на якому виділяються характерні ознаки для примітивів.
6. Етап, на якому сформований вектор, що описує примітив, піддається процедурі розпізнавання.
7. Етап, для якого виконується процедура визначення типу поверхні.

Всі етапи від 4 до 7 повторюються циклічно для всіх елементів графічного об'єкта та його проєкцій.

8. Етап на якому виконується перевірка правильності розпізнавання елементів графічного об'єкта.
9. Етап на якому можна попередньо поєднувати елементарні поверхні та відповідні примітиви, що розташовується на проєкціях. При цьому важливо відслідковувати умови об'єднання поверхонь і примітивів, а також взаємні переходи між поверхнями.
10. Етап на якому вибираються та визначаються ключові елементи, для яких власне і підлаштовуватиметься технологія виробництва.

Далі йдуть етапи, які важливі при 3D моделюванні графічного об'єкта та його синтезі, наприклад, із застосуванням технології 3D друку.

11. Етап, на якому за допомогою спеціального програмного забезпечення можна вибрати базові поверхні, які важливі при виготовленні.

- 12.Етап, на якому можна задати параметри якості для ключових поверхонь.  
Також на цьому етапі визначається і якість деталі загалом.
- 13.Етап на якому задають допуски елементів графічного об'єкта.
- 14.Етап на якому визначають матеріал з якого виготовляють графічний об'єкт та масогабаритні характеристики.
- 15.Заключний етап на якому можна безпосередньо переходити до формування технологічної карти виготовлення графічного об'єкта та застосування технологій 3D друку.

Виходячи з вище сказаного, та враховуючи потенціал застосування апарату нейронних мереж для вирішення завдань дослідження, можна узагальнити модель інтеграції задачі розпізнавання креслень до автоматизованої системи підготовки виробництва у вигляді наступної схеми, див. рис. 1.3.

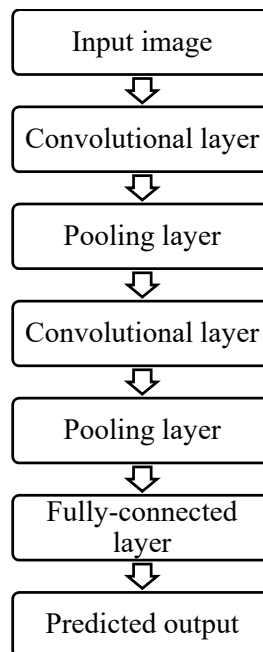


**Рисунок 1.3 Модель автоматизованої підготовки виробництва на основі розв'язання задач розпізнавання креслень та їх трансформації 2D→3D**

*Джерело: побудовано автором*

Для подальшого розвитку цих методів, моделей та інформаційних технологій, задіяних на різних етапах автоматизації проектно-конструкторських та проектно-технологічних робіт слід ширше застосовувати високотехнологічні підходи, зокрема, що базуються на теорії та методах розпізнавання образів та засобах штучного інтелекту.

Кілька публікацій [52, 53] стверджують, що можна створити 3D-модель, використовуючи єдине вхідне зображення та згортокові нейронні мережі (CNN) (рис. 1.4). Можна припустити, що його можна вважати досить простим для третьої зручної для споживача вимоги. Припущенням, яке потребує тестування, є обмеження вихідної роздільної здатності та швидкості генерації, що ніколи не було зазначено в жодній із згаданих вище документів. Основною метою даної роботи є визначення можливих обмежень використання згорткових нейронних мереж для генерації 3D моделей з урахуванням вихідної роздільної здатності та швидкості генерації. Крім того, ця робота також зосереджена на балансуванні між підвищенням якості роздільної здатності та витратами часу на генерацію вихідних даних.



**Рисунок 1.4. Алгоритм згорткової нейронної мережі**

*Джерело: [54]*

Базуючись результатах досліджень в [52, 53], спочатку береться набір двовимірних вхідних зображень, знятих під різними кутами, і кожне зображення пропускається через 2D CNN, щоб отримати представлення високого рівня на рівні презентації.

Припускають, що  $x^v \in R^{H \times W \times D}$  і  $x_{cross}^v = f_{cnn}(x^v) \in R^{D_{cross}}$  - це вхідне зображення та вивчені особливості перед шаром СС за допомогою згорткової нейронної мережі із презентації  $v$ -го відповідно, де  $H, W$  і  $D$  визначають висоту, ширина та канал. Для  $v$ -ї презентації було визначено набір  $S_v$ , який містить різні пари презентацій щодо  $v$ -ї презентації, в результаті чого

$$S_v \left| M_{v=1} = \{(v, \bar{v})\}_{\bar{v}=\frac{\{1, \dots, M\}}{v}} \right| M_{v=1} \quad (1.1)$$

де  $M$  — кількість двовимірних вхідних зображень і  $(v, \bar{v}) = (\bar{v}, v)$ . Тому запропоновану інформацію про презентацію для  $v$ -ї презентації  $x_{cross}^v$  для всіх презентацій можна обчислити за допомогою зовнішнього продукту, об'єднання пар презентацій і одновимірної згортки. Описується як:

$$x_{en}^{v, \bar{v}} = x_{cross}^v \otimes x_{cross}^{\bar{v}} \quad (1.2)$$

$$x_{en}^{S_v} = \{x_{en}^{v, \bar{v}}\}_{\bar{v}=\frac{\{1, \dots, M\}}{v}} \quad (1.3)$$

$$x_{cross}^v = f_{conv}(\Gamma(x_{en}^{S_v})) \quad (1.4)$$

де  $x_{en}^{v, \bar{v}} \in R^{D_{cross} \times D_{cross}}$  визначає зовнішній продукт презентаційної пари  $(v, \bar{v})$ , який фіксує кореляції шляхом множення кожного елемента  $x_{cross}^v$  на кожен елемент  $x_{cross}^{\bar{v}}$ . Поширюючись на всі пари презентацій  $v$ -ї презентації,  $x_{en}^{S_v} \in R^{D_{cross} \times D_{cross} \times (M-1)}$  зберігає інформацію про кореляцію  $M - 1$  презентації щодо іншої  $v$  презентації. Більше того,  $\Gamma(x_{en}^{S_v}) \in R^{D_{cross} \times D_{cross}}$  максимізує кореляції пар-презентації  $M - 1$  у  $S_v$  уздовж

напрямку різних пар презентації для презентації  $v$ , де  $\Gamma$ -це операція об'єднання пари презентації.

Інформація про перехресну презентацію рівня  $x_{cross}^v \in R^{D_{cross}}$  генерується шляхом застосування  $f_{conv}$  над  $\Gamma(x_{en}^v)$ , що складається з двох етапів, що перетворює кожен ряд  $\Gamma(x_{en}^v)$  у вектор  $k$ , застосовуючи одновимірну згорнуту (з розміром ядра = 1) і об'єднання  $D_{cross}$  векторів  $K$ -вимірності для проєктування у вектор  $D_{cross}(x_{cross}^v)$  через шар СС.

Після цього  $xv\ intra$  та  $xv\ inter$  були об'єднані за допомогою операції злиття, а потім використані як вхідні параметри для рівня СС, щоб отримати відповідні втрати та передбачення міток для кожної презентації. Описується як,

$$x_{con}^v = f_{cat}(x_{cross}^v, x^v) \quad (1.5)$$

$$z^v = f_{cc}(x_{con}^v) \quad (1.6)$$

Якщо  $x_{con}^v \in R^{D_{cross}+D}$  визначає всебічну інформацію кожної презентації, а  $z^v \in R^{N_c}$  виробляється за допомогою  $f_{cc}$  з входом  $x_{con}^v$ , що вказує на розподіл ймовірностей над можливими класами для кожної презентації, а  $N_c$  - кількість категорій. Після цього було запропоновано новий метод злиття втрат з адаптивним зважуванням із відповідним мінімумом для кількох прогнозів  $z^v | M_{v=1}$  для створення спільного рішення та реалізації класифікації 3D-моделі з кількома презентаціями, яка може бути показана як,

$$\alpha T1 = 1, \alpha \geq 0, \|\alpha\|_0 = s \sum_{v=1}^M \alpha_v^y L^v(z^v, y) \quad (1.7)$$

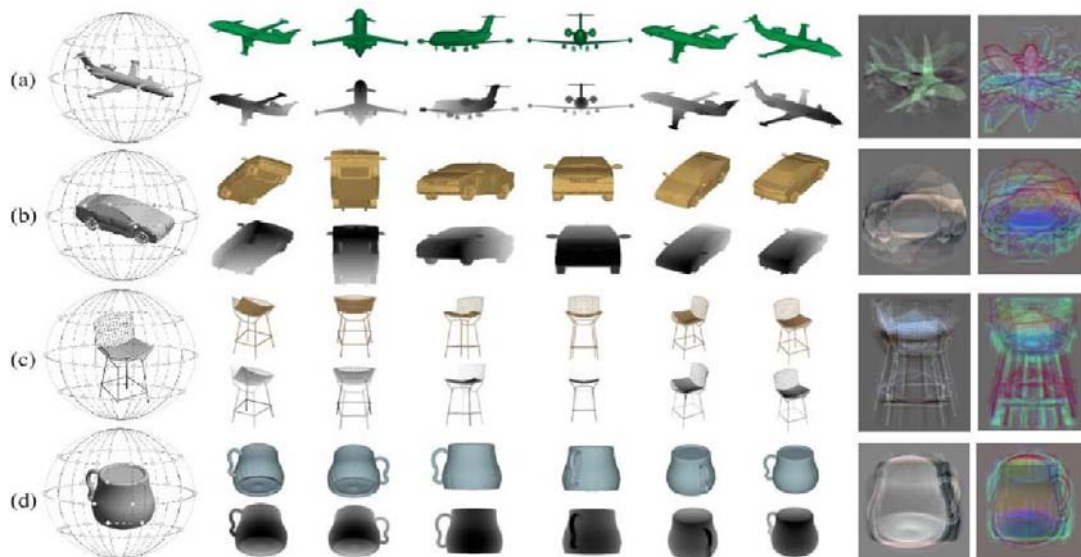
$$L^v(z^v, y) = -\log\left(\frac{\exp(z_y^v)}{\sum_{o=1}^{N_c} \exp(z_o^v)}\right) \quad (1.8)$$

де  $\alpha \in R^M$  - ваговий вектор, що відповідає декільком презентаціям,  $y \in R$  визначає загальну інформацію мітки всіх презентацій для об'єкта, а  $L^v(z^v, y) \in R$  є втратою крос-ентропії  $v$ -ї презентації.  $\gamma > 1$  є параметром степеневого показника ваги  $\alpha_v$ , який гнучко регулює розподіл ваги різних видів і дозволяє уникнути тривіального рішення  $\alpha$  під час класифікації.  $\|\alpha\|_0 = s$  використовується для обмеження розрідженості вектора ваги  $\alpha$ , де  $s \in N_+$  позначає кількість ненульових елементів у  $\alpha$ . Важливо те, що обмеження  $Z_0$ -норма здатне охопити глобальні зв'язки між різними видами та може досягти розрідженості представлення, так що лише кілька дискримінаційних та інформативних переглядів вибираються під час оптимізації для прийняття рішень

Після цього запропонований метод на наборі даних Model Net 40 було оцінено та проведено порівняння з кількома сучасними методами. Як це зазвичай буває, класифікація в 3D в основному базується на моделі автоматизованого проєктування (CAD). Лише один широко використовуваний репозиторій Model Net [55] містить понад 130 тисяч 3D-моделей CAD із понад 600 категорій. Model Net 40 [56], представлений на веб-сайті Princeton Model Net (рис.1.4.), є підмножиною Model Net і містить близько 12 тисяч моделей із 40 загальних категорій. На рисунку 1 вибрано 4 типи простих категорій, щоб інтуїтивно показати шість 2D-презентацій, візуалізованих з 3D-моделей, де шість презентацій генеруються з 360 градусів з інтервалом 60 градусів (для самого експерименту використовували 12 презентацій тих самих моделей з інтервалом 30 градусів).

Запропонований метод порівнювався з кількома найсучаснішими методами класифікації мультипрезентаційних 3D-моделей, включаючи методи презентації та об'єму MVCNN Multi Res, Minto, загальні об'ємні методи 3DShapeNets, загальні точкові методи PointNet та загальну презентацію на основі методів MVCNN.





**Рисунок 1.4: Приклади презентацій, створених за категоріями 3D моделей у Model Net 40**

*Джерело: [54]*

З огляду на FLOP глибшої згорткової нейронної мережі та кращий компроміс між точністю та рівнем споживання пам'яті порівняно з іншими класичними згортковими нейронними мережами Res Net 18 є хорошим вибором, але не обмежується цією архітектурою згорткової нейронної мережі. Щоб оцінити базові архітектури, результати Multi View Convolutional Neural Network порівнювалися з ResNet18, результати яких наведено в таблиці 1. Очевидно, що використання Res Net 18 може призвести до підвищення продуктивності Multi View Convolutional Neural Network. Наприклад, Multi View Convolutional Neural Network (ResNet18) із 12 переглядами досягає 3,3% покращень порівняно з автономною Multi View Convolutional Neural Network. Для запропонованого методу параметри Resnet-18 були відкалібровані за допомогою набору даних Model Net 40 та використовують ADAM з  $learning\ rate = 5 * 10 - 6$ ,  $\beta_1 = 0,7$ ,  $\beta_2 = 0,933$ ,  $weight\ decay = 0,0001$ ,  $batch = 8$ ,  $epoch = 60$  для оптимізації. Крім того, у запропонованому методі є два параметри  $s$  і  $\gamma$ , де  $s$  представляє кількість ненульових елементів у  $\alpha$ , а  $\gamma$  є показником степеня кожного елемента  $\alpha$ .  $s$  було відкалібровано в діапазоні [57, 58] з кроком 1 для вибору

кількох дискримінаційних та інформативних презентацій для створення спільного рішення під час класифікації.  $\gamma$  змінювався від 1,5 to 10 з кроком 1, щоб дослідити вплив різних значень  $\gamma$  на точність класифікації. На основі відповідних параметрів  $s = 9$  і  $\gamma = 2,5$  можна навчити оптимальну модель, щоб значно покращити продуктивність класифікації 3D-моделей (див. таблицю 1.4).

Таблиця 1.4

Результати точності з використанням існуючих методів і запропонованого способу

Метод	Точність
MVCNN	89.90%
MVCNN (ResNet18)	92.20%
Proposed (ResNet18)	93.01%

Джерело: [54]

Продуктивність різних модулів запропонованого методу була оцінена та наведена в таблиці 1.5. Продемонстровані дослідження показують, що розподіл ваги  $\alpha$ , розрідженість кількох презентацій та інформація про перехресне представлення для будь-яких різних презентацій відіграють різні ролі під час класифікації 3D-моделей.

Таблиця 1.5

Точність результатів при використанні різних підходів до запропонованого способу

Метод	Точність
Proposed (merge)	88.02%
Proposed ( $\alpha$ )	92.61%
Proposed ( $s$ )	92.23%
Proposed ( $\alpha + s$ )	92.98%
Proposed ( $\alpha + s + cross$ )	93.01%

Джерело: [54]

На підставі вищесказаного можна конкретизувати завдання дослідження, які вирішуватимуться в наступних розділах дисертації:

1. Обґрунтувати вибір конкретної моделі згорткової нейронної мережі на дослідження у межах цілей дисертації.
2. Описати математичну модель згорткової нейронної мережі завдання розпізнавання типових примітивів на кресленнях, з прикладу типових позначень.
3. Розглянути вимоги до навчальної вибірки згорткової нейронної мережі.
4. Проаналізувати потенціал покращення роботи згорткової нейронної мережі, а також її архітектури, не вдаючись до екстенсивного розширення архітектури згорткової нейронної мережі, а також враховуючи необхідність вирішення завдання, пов'язаного з логічною векторизацією розпізнаних за допомогою згорткової нейронної мережі примітивів та/або умовно-графічних позначень на кресленнях для виконання трансформації  $2D \rightarrow 3D$ .
5. Розробити концептуальну модель інформаційної технології розпізнавання графічної інформації на основі нейронної мережі.
6. Описати генезу розпізнавання графічної інформації на основі нейронної мережі.
7. Розробити декомпозиції моделей нотації IDEF0 та IDEF3 засобами CASE-технології ERwin.
8. Спроекувати архітектуру програмної реалізації інформаційної технології розпізнавання графічної інформації на основі нейронної мережі.

## ВИСНОВКИ ДО РОЗДІЛУ 1

Таким чином, в результаті досліджень, виконаних у першому розділі роботи, були отримані такі результати:

- Виконано огляд та аналіз попередніх досліджень та підходів у питаннях розпізнавання графічної інформації з креслярсько-технічної документації та для подальшої трансформації моделей  $2D$  у  $3D$ . При цьому встановлено, що проблематика розпізнавання креслень та трансформації  $2D \rightarrow 3D$  є дуже актуальною в умовах розвитку різних ІТ та прикладного програмного забезпечення для системи автоматизованого проєктування, зокрема, орієнтованих на використання  $3D$  моделей у ході вирішення прикладних технічних, технологічних, наукових та інших завдань.
- Встановлено, що науково-технічне завдання, пов'язане з розпізнаванням креслень та трансформацією  $2D \rightarrow 3D$  на основі проєкційних зображень, залишається не до кінця вирішеним. Причому потребують розвитку як її алгоритмічна складова, так і програмно-практична частина, орієнтована на створення програмних продуктів.
- Розглянуто потенціал застосування різних методів штучного інтелекту в задачі розпізнавання креслень та трансформації  $2D \rightarrow 3D$ . Причому, як показав аналіз літературних джерел, досі існує як практичний, так і науковий інтерес до цієї тематики, що й зумовлює велику кількість публікацій, пов'язаних із цим напрямком.
- Показано, що проблематика трансформації  $2D \rightarrow 3D$  може бути ефективно вирішена на основі застосування систем штучного інтелекту, зокрема на основі застосування штучних нейронних мереж, що дозволить значно знизити трудомісткість вирішення окремих підзавдань як в алгоритмічному, так і в практичному аспектах.

- Систематизовано етапи трансформації 2D зображень у 3D модель в класичних підходах.

Основні результати розділу висвітлені у науково-дослідній роботі НДР № 0122U001549 «Моделювання інтелектуальних систем управління діяльністю підприємств», (довідка від 29.08.2023 №1453/24)

Основні результати розділу опубліковані в наукових працях автора:

Наукові праці, в яких опубліковані основні наукові результати дисертації:

[2], [3], [4], [5], [8].

Наукові праці, які засвідчують апробацію матеріалів дисертації:

[10], [11], [12], [13].

Наукові праці, які додатково відображають наукові результати дисертації

[14], [15], [16], [17], [18], [19], [20], [21], [22].

## РОЗДІЛ 2

### ЗАСТОСУВАННЯ ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ З МОДУЛЕМ ЕЛЕМЕНТАРНИХ КЛАСИФІКАТОРІВ ГРАФІЧНИХ ПРИМІТИВІВ У ЗАДАЧАХ РОЗПІЗНАННЯ КРЕСЛЕНЬ І ТРАНСФОРМАЦІЇ МОДЕЛЕЙ 2D В 3D

Як було встановлено за результатами огляду та аналізу попередніх досліджень (див. розділ 1), серед усіх моделей та методів штучного інтелекту (ШІ), що використовуються для розпізнавання образів, найбільш перспективними є штучні нейронні мережі (ШНМ).

Як було показано в роботах [24, 26, 43-45, 59-62] застосування класичних штучних нейронних мереж у задачі розпізнавання графічної інформації, у тому числі креслярської документації, не дасть належного ефекту з наступних причин:

- зображення, що розпізнаються на креслярській документації, як правило, характеризуються великою розмірністю, в свою чергу зростає розмір нейронної мережі;
- існує необхідність мати досить велику тренувальну вибірку, а це складно реалізувати на практиці з огляду на величезну різноманітність номенклатури креслярської документації. Крім того, операції з такими великими масивами даних призводять до зростання часу на навчання і сама процедура навчання ускладнюється в міру ускладнення зображення проєкцій деталей на кресленнях;
- використання двох і більше нейронних мереж [24, 26] на різних етапах розв'язання задачі розпізнавання креслярської документації, хоч і дає позитивний ефект, але радикально не скорочує сумарний час розв'язання, а також не призводить до суттєвого зниження часу розв'язання;

- класичні штучні нейронні мережі, як показав ряд прикладних досліджень [24, 52, 54], не дають прийнятних результатів у разі інваріантності зображень, а також при зміні масштабу зображення.

На підставі перерахованого вище, для вирішення задачі розпізнавання креслень і трансформації  $2D \rightarrow 3D$  були обрані згорткові нейронні мережі.

## **2.1 Архітектура згорткової нейронної мережі з модулем елементарних класифікаторів графічних примітивів для вирішення задачі розпізнавання креслень та трансформації $2D \rightarrow 3D$**

Кожен шар згорткової нейронної мережі [62-68] умовно можна уявити, як набір площин (див. рис. 2.1). У свою чергу кожна площина складатиметься з нейронів. Нейрони відповідної площини характеризуються однаковими синаптичними коефіцієнтами. Кожен із нейронів, що знаходиться у відповідному шарі, отримуватиме вхідне зображення попереднього шару. CNN представляється як багат шаровий персептрон [62].

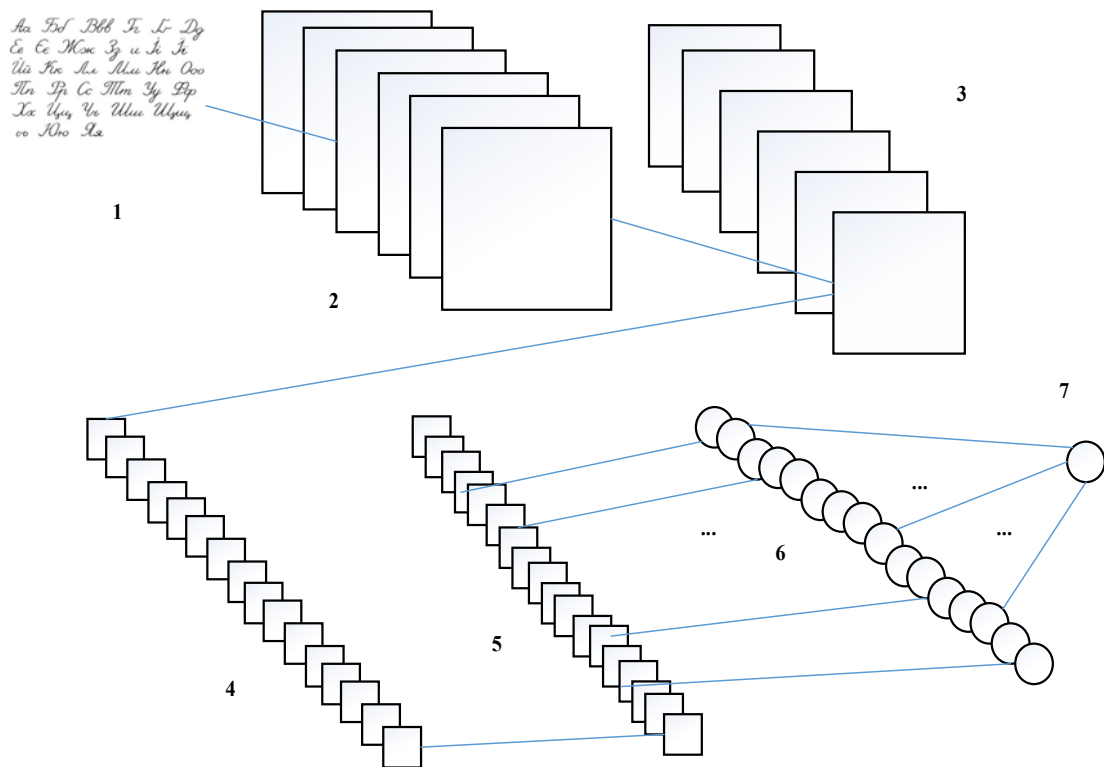
Розглянемо етап вилучення ознак. Кожен нейрон отримує вхідні сигнали від локального рецептивного поля, що знаходиться в попередньому шарі. На цьому етапі реалізується процедура вилучення обмежених ознак. Як тільки виконано вилучення ознаки її наступне розташування не матиме значення.

Далі розглянемо етап відображення ознак. Як показано на рисунку 2.1, обчислювальні шари мережі CNN включають множини т.зв. карт ознак [62-64].

Кожна карта ознак може бути представлена як площина. На цій площині нейрони спільно використовують свою множину синаптичних ваг [65].

Далі розглянемо етап підвибірки. За відповідними шарами згортки на рисунку 2.1 показані обчислювальні шари. Кожен такий обчислювальний шар підходить для операції локального усереднення та підвибірки [65, 66].

Локальне усереднення дозволяє досягти скорочення дозволу для карт ознак [65, 66]. Фактично це забезпечує послаблення чутливості вихідних сигналів операторів відображення ознак [66-68].



1- Вихідне зображення; 2, 3,4,5-картки ознак; 6,7-згортка

**Рисунок 2.1. Схема CNN для розпізнавання рукописного тексту (включає один вхідний, чотири приховані та один вихідний шар нейронів)**

*Джерело: побудовано автором на основі джерел [66-68]*

На рисунку 2.1 представлена схема згорткової нейронної мережі, яка включає в свою структуру один вхідний, чотири приховані і один вихідний шар нейронів. Подібна мережа підходить для розпізнавання рукописного тексту [66-68]. І хоча така задача теж важлива для розпізнавання



креслярської документації, але не орієнтована на розпізнавання креслярських примітивів. і, скоріше може бути складовою частиною при вирішенні загального завдання дослідження.

Щоб зробити розпізнавання образів, які складають зображення на кресленні або креслярських примітивах (лінії та стрілки; прямокутники; кола, еліпси, дуги, сегменти та сектори; криві; сполучні лінії; тривимірні об'єкти (куби, кулі, циліндри, тори тощо)) ; текстові дані), наприклад, як це показано на рис. 2.2, необхідний процес навчання штучної нейронної мережі.

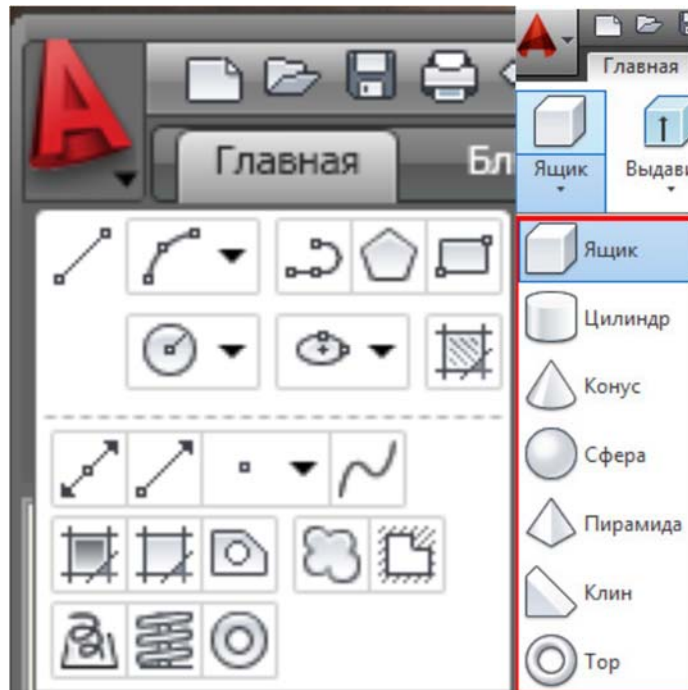


Рисунок 2.2. Приклад примітивів для формування зображення в **AutoCAD**

*Джерело: побудовано автором (скріншот з екрану)*

Для забезпечення однозначності, повноти та цілісності класифікації наявних множин примітивів, задіяних у процесах розпізнавання креслярсько-технічної документації та подальшої трансформації  $2D \rightarrow 3D$ , введемо такі вимоги до класифікації примітивів, задіяних у згаданих процесах:

- непересічні класи графічних примітивів (визначає однозначність вибору класу на основі зовнішнього правила, що дозволяє прийняти рішення);
- застосовність (додавання класу не повинно викликати дроблення більше одного класу на дві частини);
- об'єктивність (наявність чи відсутність класу має підтверджуватись відомими класифікаціями);
- розширюваність (додавання класу можливе шляхом дроблення існуючих класів);
- кількість класів злічена.

Базисом для побудови елементарних класифікаторів (ЕК) примітивів для згорткової нейронної мережі, що використовується у процесах розпізнавання креслярсько-графічної документації, можуть стати різні дані. Наприклад, описова інформація представлена у формі двійкового представлення координат точок на вихідному кресленні  $\{p_{ax1}, \dots, p_{axn}\}$ .

Прийmemo:

$MI$  – загальна кількість примітивів, що використовуються для процедури навчання та подальшого розпізнавання;

$PA$  – кількість потенційних примітивів для конкретної сфери креслярсько-технічної документації (машинобудування, будівництво та ін.);

$B_{pa}$  – множина номерів фактичних примітивів під час процедур навчання CNN.

У загальному випадку проблематика розпізнавання креслярсько-технічної документації полягає в наступному [59-68]:

- Досліджуємо множину об'єктів, у нашому випадку це  $PA$  – число потенційних примітивів (зокрема і літерно-цифрові). Об'єкти, які можна віднести до даної множини, характеризуються сукупністю ознак  $\{p_{ax1}, \dots, p_{axn}\}$ . Відомо, що множина  $PA$  представлена у вигляді об'єднання підмножин

(класів) примітивів -  $(KL_1, \dots, KL_l)$  або  $(B_{pa1}, \dots, B_{pal})$ , які не перетинаються.

- Припустимо, що існує скінченна група об'єктів  $\{sp_{a1}, \dots, sp_{am}\}$  з РА, про які відомо, до яких класів примітивів їх можна віднести. Це прецеденти, тобто, об'єкти, що використовуються для навчання (далі ОВН) згорткової нейронної мережі. CNN необхідна за наявною множиною параметрів (ознак), тобто використовуючи опис об'єкта  $sp_{an}$  із РА, класифікувати цей об'єкт. Потім, за підсумками класифікації, вибудовується робота автоматизованої або автоматичної системи розпізнавання креслярсько-технічної документації. CNN спочатку невідомо до якого класу належить об'єкт на кресленні. Це може бути деталь машини, труба, будівельна конструкція тощо.

Автоматизація пошуку нових примітивів передбачає застосування моделей, що базуються на комбінаторному аналізі описів ознак  $\{sp_{a1}, \dots, sp_{am}\}$ . У роботах багатьох учених [69-72] було показано, що ці методи дають прийнятний результат розпізнавання в ситуації, коли дані представлені в дискретній формі і кількість значень ознаки, що допускаються, невелика [69].

Винятковою особливістю запропонованих процедур імплементації ЕК для розпізнавання графічних примітивів на креслярсько-технічній документації є можливість отримати результат при навчанні згорткової нейронної мережі навіть у ситуації, коли немає даних щодо функцій, що описує розподіл значень ознак цих примітивів. Такі процедури надалі називатимемо логічними процедурами.

Ставиться завдання мінімізації навчальної вибірки для розпізнавання зображення у межах класу, наприклад, лише сфери машинобудівних креслень. При цьому немає необхідності задавати так звані метрики в

просторі описів об'єктів, що характеризують кожен клас графічних примітивів.

Парадигмою побудови бази примітивів для навчання згорткової нейронної мережі (чи логічних процедур навчання ЛПН) є пошук інформативних фрагментів описів відповідних графічних примітивів. Ці фрагменти при створенні конкретних проєктних рішень для систем розпізнавання креслярсько-будівельної документації, наприклад, на основі згорткових нейронних мереж, дозволять однозначно робити висновок про наявність (або відсутність) в базі системи розпізнавання певного графічного примітиву в рамках класу.

Інформативними покладемо фрагменти, що відображають характерні закономірності при описі об'єкта, що використовується під час навчання CNN у завданнях розпізнавання креслярсько-технічної документації. Тоді наявність (відсутність) фрагмента(ів) в описі об'єкта, що проходить класифікацію, дає можливість віднести його до певного класу. У ЛПН інформативним покладемо фрагмент(и), який є в описах об'єктів класу, що розглядається в класах графічних примітивів, але відсутній в описі об'єктів інших класів. Розглянуті фрагменти необхідно доповнити описами термінах проєктування згорткової нейронної мережі.

При побудові ЛПН для CNN використовують запропоновані рядом авторів Гуптою С., Кумаром М., Гаргом А. та Гадатом С. [73, 74] ЕК. ЕК – це фрагмент, який описує об'єкт. У свою чергу, цей об'єкт використовується для навчання CNN у завданнях розпізнавання креслярсько-технічної документації. Для кожного класу зображень  $(KL_1, \dots, KL_l)$  виконується побудова множини ЕК із заздалегідь заданими параметрами.

При цьому наступні міркування для елементарних класифікаторів засновані на такій гіпотезі, що складається з двох припущень:

- 1) використовуються елементарні класифікатори, які присутні в описах об'єктів класу, що аналізується в даний момент (наприклад, машинобудівних креслень), але їх немає в описі об'єктів інших класів;

2) описовий набір показників (ознак) для відповідного графічного примітиву задається у двійковій формі (наприклад, 0010101). При цьому об'єкти, що використовуються для навчання характеризують усі об'єкти цього класу. А отже, ОВН мають більшу інформативність. Для підвищення ефективності логічних процедур навчання актуальною є проблематика використання в них властивості «незустрічності» груп із прийнятних значень показників графічних примітивів. Крім того, необхідно імплементувати в алгоритм розпізнавання «вирішальне правило»  $DR_{(p_{axi})}$  для CNN з мінімальною кількістю помилок у його роботі.

Наступна проблема при проєктуванні згорткової нейронної мережі - присутність у вибірці об'єктів, що використовуються для навчання, з характеристиками, що лежать на стику різних класів графічних примітивів, наприклад, квадрат-прямокутник або коло-овал. Подібний об'єкт, що використовується для навчання є атиповим для свого класу. Це пояснюється схожістю інформативних підписань, так як показники, подані у двійковій формі в цілому близькі. Присутність у нейронній мережі атипових ОВН, збільшує довжину інформативних підписів графічних примітивів. Таким чином, можна буде розрізняти об'єкти з різних класів. А оскільки довгі інформативні підписи не часто можна зустріти для нових об'єктів, це призводить до збільшення кількості нерозпізнаних елементів графічно-креслярської документації.

Синтез алгоритмів ефективних реалізацій ЛПН для CNN пов'язані з результатом аналізу кількісної оцінки інформативності множин фрагментів, тобто, показників графічних примітивів.

При вирішенні завдань, пов'язаних з проєктуванням ефективних згорткових нейронних мереж, коректні дані про структуру всього об'єкта розпізнавання в деяких ситуаціях можуть бути відсутніми. Отже, спочатку, при синтезі ЛПН не гарантується коректність їх використання на нових об'єктах, які відмінні від  $\{sp_{a1}, \dots, sp_{am}\}$ .

Іншими словами, ті примітиви, які пройшли навчання для машинобудівних креслень, не працюватимуть для будівельної сфери, хоча візуальна схожість може бути присутньою. При цьому, якщо об'єкти, що використовуються для навчання мають характеристики властиві для аналізованої множини  $\{sp_{a1}, \dots, sp_{am}\}$ , то алгоритм, який працює безпомилково на етапі навчання згорткової нейронної мережі, забезпечить прийнятні результати і на некласифікованих  $sp_{an}$ , які не входили у вибірку об'єктів, що використовуються для навчання. У зв'язку з цим, у роботах таких дослідників як Гадат С., Фроб Б., Ернст А., ЛеКун Ю., Хаффнер П., Ботту Л. та Бенгіо Ю. [74-76] приділялося багато уваги проблематиці коректності алгоритмів, що використовуються для розпізнавання за допомогою CNN. Алгоритм коректний, якщо він чітко розпізнає об'єкти з тестової вибірки.

Працездатність алгоритму може бути перевірена за допомогою найпростішої процедури. Аналізований об'єкт  $sp_{an}$ , наприклад, описові показники (ОП) графічного примітиву тощо, порівнюється з кожним з ОВН  $\{sp_{a1}, \dots, sp_{am}\}$ .

У випадку якщо описові показники об'єкта  $sp_{an}$  тотожно з описовими показниками об'єкта, що використовується для навчання  $sp_{ai}$ , його слід віднести до того класу, до якого належить  $sp_{ai}$ . Інакше алгоритмом буде видано повідомлення про відмову від розпізнавання. Алгоритм працездатний [73], але він не адаптований розпізнати жоден об'єкт  $sp_{ai}$ , ОП якого не тотожні з описовими показниками об'єкта, що використовується для навчання, що наявні у репозиторії [74-77].

В рамках даного підрозділу дисертації пропонується більш зручний спосіб вирішення одного з локальних завдань розпізнавання креслярсько-технічної документації за допомогою CNN. При цьому можна виконати перехід від вихідного простору описів об'єктів  $P_{ax_j}$  у формі двійкового представлення ознак до нового простору  $NEP_{ax}$  за допомогою перетворення  $P_{ax_j} \rightarrow NEP_{ax}$  [67-70].

З урахуванням робіт [67-70] приймемо, що  $NP_{p_a}$  - деяка добірка з  $r_{p_a}$ ,  $r_{p_a} \leq MI$  неподібних дискретних показників (наприклад, координати точок вихідного зображення, тип ліній (суцільна, пунктирна, штрихова, тощо) виду  $\{p_{aj_1}, \dots, p_{aj_r}\}$ . Відстань між об'єктами  $sp'_a = (\alpha p'_{a1}, \alpha p'_{a2}, \dots, \alpha p'_{aMI})$  та  $sp''_a = (\alpha p''_{a1}, \alpha p''_{a2}, \dots, \alpha p''_{aMI})$  з РА за набором  $NP_{p_a}$  оцінимо величиною

$$BN(sp'_a, sp''_a, NP_{p_a}) = \begin{cases} 1, \text{ якщо } \alpha p'_{j_{ti}} = \alpha p''_{j_{ti}} \text{ при } ti = 1, 2, \dots, r_{p_a} \\ 0 \text{ в іншому випадку} \end{cases} \quad (2.1)$$

У системі показників ОВН  $\{p_{a_1}, \dots, p_{a_{jMI}}\}$  виділимо сукупність підмножин виду:

$$NP_{p_a} = \{p_{aj_1}, \dots, p_{aj_{MI}}\}, r_{p_a} \leq MI.$$

Виділені підмножини вважатимемо опорними множинами (ОМ) алгоритму розрахунку оцінок (АРО). Всю їхню сукупність позначимо –  $\Omega MI$ .

Задамо такі додаткові параметри:

$po_{sp_a}$  - параметр, що характеризує значимість об'єкта  $sp_{ai}$ ,  $i = 1, 2, \dots, PA$ ;

$po_{NP_{p_a}}$  - параметр, що характеризує значимість об'єкта ОМ  $NP_{p_a} \in \Omega MI$ .

Виконаємо розрахунок оцінки, порівнюючи об'єкт  $sp_{an}$  із кожним ОВН для CNN  $sp_{ai}$  по кожній ОМ.

Для кожного класу графічних примітивів, що використовуються для навчання CNN  $KL$ ,  $KL \in \{KL_1, \dots, KL_l\}$ , обчислимо оцінку належності  $\Gamma(sp_a, KL)$  об'єкта  $sp_a$  класу  $KL$ :

$$\Gamma(sp_a, KL) = \frac{1}{|LW_{KL}|} \sum_{sp_{ai} \in KL} \sum_{NP_{p_a} \in \Omega MI} po_{sp_a} \cdot po_{NP_{p_a}} \cdot BN(sp_a, sp_{ai}, NP_{p_a}), \quad (2.2)$$

де  $|LW_{KL}| = |KL \cap \{sp_{a1}, \dots, sp_{aMI}\}|$ .

Об'єкт  $sp_{an}$  відноситься до класу, що має найбільшу оцінку. Якщо є множина таких класів, то алгоритм відмовляється від подальшого розпізнавання. Для підвищення коректності алгоритму необхідно вирішити таку систему нерівностей:

$$\begin{aligned} \Gamma(sp_{a1}, KL_1) &> \Gamma(sp_{a1}, KL_2) \\ \Gamma(sp_{aMI_1}, KL_1) &> \Gamma(sp_{aMI_1}, KL_2) \\ \Gamma(sp_{aMI}, KL_2) &\overset{\dots}{>} \Gamma(sp_{aMI}, KL_1) \end{aligned} \quad (2.3)$$

Для розв'язання системи (2.3) необхідно вибрати параметри

$$po_{sp_{ai}} \quad i = 1, 2, \dots, PA \text{ та } po_{NP_{pa}}, NP_{pa} \in \Omega MI.$$

Коли система несумісна, необхідно знайти для неї максимальну сумісну підсистему. Потім із рішення цієї підсистеми визначити значення  $po_{sp_{ai}}$  і  $po_{NP_{pa}}$ .

Альтернативним способом підвищити коректність роботи алгоритму є шлях добірки системи достовірних опорних множин для того, щоб розпізнати та класифікувати графічні примітиви на креслярсько-технічній документації.

Наприклад, вибрати добірку таким чином, щоб для будь-якого ОВН  $sp'_a \notin KL$  виконувалася умова  $\Gamma(sp'_a, KL) = 0$ . Крім того, для будь-якого об'єкту, що використовується для навчання  $sp''_a \in KL$  виконувалася нерівність  $\Gamma(sp''_a, KL) > 0$ .

Реалізувати це можна в такий спосіб.

$$\text{Припустимо, що } NP_{pa} = \{p_{aj_1}, \dots, p_{aMI}\} \in \text{ОМ}.$$

Вважатимемо, що сукупність показників (або ознак)  $NP_{pa}$  задовольняє вимогам тесту, якщо для кожного об'єкта, що використовується для навчання  $sp'_a$ ,  $sp''_a$ , і при цьому, що належать до несхожих класів, виконується умова

$$BN(sp'_a, sp''_a, NP_{pa}) = 0 \quad (2.4)$$



Таким чином, наш тест – це група (підбірка) показників, за яким розрізняються лише будь-які об'єкти з різних класів.

Враховуючи попередні викладки, можна стверджувати, що на вхід згорткової нейронної мережі будуть подаватися деякі зображення (оцифровані дані в двійковій формі – закодовані зображення). У свою чергу, на виході CNN маємо необхідний вид зображення на підставі трактування мережею простої відповідності набору класу/класів об'єкта.

Отже, можна дійти певного висновку. Вхідний сигнал можна трактувати як постійну умовну величину. Вихідний сигнал буде являти собою дискретне значення. При такій комбінації застосування апарату елементарного класифікатора і безпосередньо CNN утворюється перевага порівняно із звичайною процедурою застосування нейронних мереж для розпізнавання. А саме виникає властивість адаптивності мережі, яка після навчання зможе працювати з різними класами об'єктів розпізнавання, як показано на рисунку 2.3.

Завдяки властивості адаптивності подібна згорткова нейронна мережа, на наш погляд, зможе досить ефективно виконувати розпізнавання різних графічних зображень, що належать до різних типів графічних примітивів.

У запропонованій архітектурі згорткової нейронної мережі для вирішення задачі розпізнавання креслень та трансформації  $2D \rightarrow 3D$ , (див. рис. 2.3) присутні:

- модуль для синтезу гіпотези про належність об'єкта до певного класу графічних примітивів. Даний модуль призначений для синтезу гіпотези/гіпотез, які відповідають за визначення координат конкретного об'єкта на зображенні, що важливо для подальшого отримання наборів показників (ознак) для відповідного графічного примітиву задається;
- модуль з елементарними класифікаторами, який власне і формує згорткову нейронну мережу.

Що стосується архітектури та безпосереднього досвіду застосування CNN для вирішення завдань розпізнавання зображень, то цьому питанню присвячено досить велику кількість раніше згаданих робіт, наприклад, [64-69].

У зазначених роботах, а також, враховуючи полеміку висновків, зроблених у роботах [78-81], можна констатувати, що в міру розвитку технологій розпізнавання та зростання обчислювальних потужностей комп'ютерної техніки стало можливим не лише підвищити ефективність систем розпізнавання на базі згорткової нейронної мережі, але й навчати такі мережі більш результативно. Причому результативність такого навчання безпосередньо залежить від успішного розв'язання задач класифікації зображень.

Вищесказане і зумовлює релевантність вирішення наступного завдання - побудова моделі класифікатора графічних зображень на основі покриттів класів та елементарних класифікаторів примітивів для підвищення ефективності навчання згорткової нейронної мережі.

Нижче викладено основні правила конструювання ЕК розпізнавання графічних примітивів на креслярсько-технічній документації.

Дані елементарного класифікатора можуть становити основу модуля ЕК, який показано на рисунку 2.3.

Поставимо у відповідність елементарного класифікатора об'єкта:

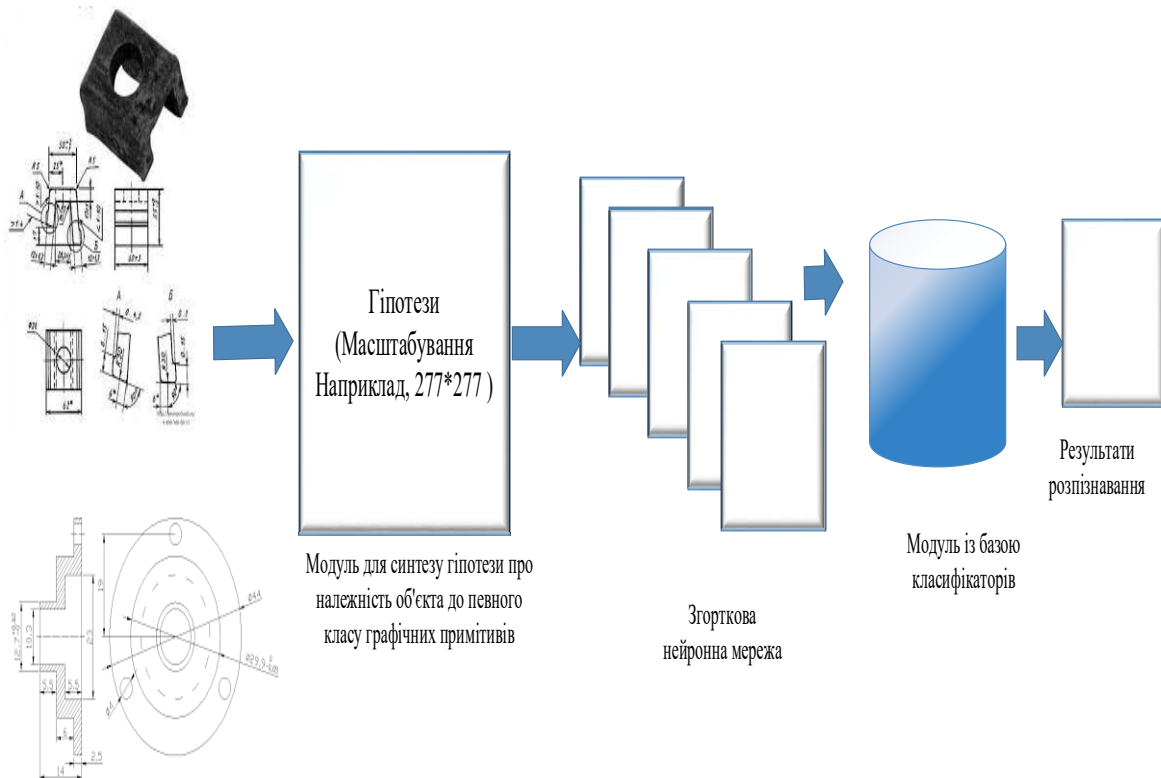
$$(\sigma_{DOP}, NP_{p_a}), \quad (2.5)$$

$$\text{де } \sigma_{DOP} = (\sigma_{DOP_1}, \dots, \sigma_{DOP_r}),$$

$NP_{p_a}$  – набір показників графічних примітивів на креслярсько-технічній документації з номерами  $j_1, \dots, j_{r_{pa}}$  елементарну кон'юнкцію

$$\mathfrak{R} = p_{axj_1}^{\sigma_{DOP_1}} \cdot \dots \cdot p_{axj_{r_{pa}}}^{\sigma_{DOP_{r_{pa}}}}.$$

Якщо  $sp_a = (\alpha p_{a1}, \dots, \alpha p_{aMI})$  - об'єкт з множини  $PA$ , то, очевидно,  $BN(\sigma_{DOP}, sp_a, NP_{pa}) = 1$  тоді і тільки тоді, коли  $(\alpha p_{a1}, \dots, \alpha p_{aMI}) \in NI_{\mathfrak{A}}$  де  $NI_{\mathfrak{A}}$  - інтервал істинності елементарної кон'юнкції  $\mathfrak{A}$ .



**Рисунок 2.3. Принципова схема CNN із модулем ЕК графічних примітивів для розпізнавання креслярсько-технічної документації**

*Джерело: побудовано автором*

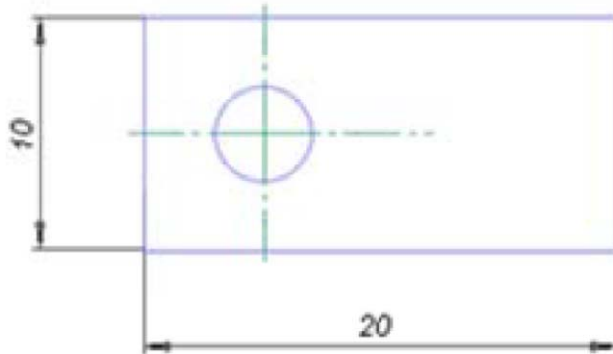
Для графічного відображення креслярських примітивів у такій популярній системі креслення, як Autodesk AutoCAD використовується технологія ActiveX. Як альтернативний варіант виступає можливість використовувати вбудовану мову програмування AutoLISP. Синтаксис останнього ближче до ідеї застосування ЕК, оскільки дозволяє безпосередньо працювати з масивами точок, що визначають координати двовимірної проєкції, наприклад, як показано на рисунку 2.4.

```

1 (SETQ pt1 (LIST 0.0 0.0))
2 (SETQ pt2 (LIST 0.0 10.0))
3 (SETQ pt3 (LIST 20.0 10.0))
4 (SETQ pt4 (LIST 20.0 0.0))
5 (COMMAND "LINE" pt1 pt2 pt3 pt4 "Close")
6 (COMMAND "CIRCLE" (LIST 5.0 5.0) 2.0)
7

```

Приклад коду в AutoLISP  
для опису координат  
вершин виробу



Проекція виробу

**Рисунок 2.4. Приклад коду в AutoLISP для опису координат вершин графічного об'єкта.**

*Джерело: побудовано автором*

Таким чином, по суті вже в середовищі Autodesk AutoCAD спочатку присутня можливість опису графічних об'єктів у вигляді числової інформації, що відповідає вершинам і ребрам графічного об'єкта, що робить завдання синтезу елементарних класифікаторів практично значущою. У будь-якому випадку після розпізнавання проєкцій відповідного графічного об'єкта за наявними кресленнями ці дані будуть вихідною інформацією для формування електронних креслень AutoCAD. Примітивами рисунку 2.4 виступають зображення прямокутника з розмірами 20\*10 мм, і навіть кола діаметром 5 мм. Повне зображення аналізованої проєкції деталі може бути охарактеризовано набором параметрів об'єкта в цифровій формі (прямокутник 10\*20) & (коло з радіусом 5). Таким чином, операція сформована за допомогою операції кон'юнкції.

Взаємне розташування креслярських примітивів може бути описано на основі матриць суміжності. Так, наприклад, стовпці описуватимуть дуги чи кола, а рядки – відрізки прямих. У такому випадку на перетинах рядків і стовпців будуть розміщені цифри, які визначають суміжність креслярських

примітивів. У модулі елементарного класифікатора, який був раніше показаний на рисунку 2.3, можна зберігати всі матриці, які визначають шаблон графічних об'єктів певної категорії.

При проектуванні згорткової нейронної мережі з модулем елементарного класифікатора графічних примітивів для розпізнавання креслярсько-технічної документації, слід врахувати, що визначення множини ЕК зводиться до відшукування допустимих і максимальних кон'юнкцій для характеристичної функції класу об'єкта  $(KL_l) = (B_{p_{al}})$ . До класів примітивів на креслярсько-технічній документації можна віднести прості зображення, показані на рисунку 2.2.

Причому, дана функція є двозначною логічною функцією, що приймає різні значення на ОВН з  $KL_l$  та  $\overline{KL_l}$  [77].

Характеристична функція класу  $KL_l$  – буде представлена у вигляді функції алгебри логіки (булевої функції)  $F_{KL_l}$ , яка дорівнює нулю (0) на інформаційних описах об'єкта  $sp_{an} = (\alpha p_{an1}, \dots, \alpha p_{anMI})$  з  $KL_l$  і дорівнює одиниці (1) на наборах показників, що залишилися з  $E_{KL}^{MI}$ .

Тут  $E_{KL}^{MI}$  - сукупність множин всіх наборів, що мають довжину  $r_{p_a}$ . У цьому випадку, покриттю класу  $KL_l$  ставиться у відповідність допустима для  $F_{\overline{KL}}$  кон'юнкція. Суцільному покриттю буде відповідати - максимальна для  $F_{\overline{KL}}$  кон'юнкція. Допустима  $\mathfrak{R}$  в матрицях показників об'єктів визначить належність конкретного об'єкта  $sp_{an} = (\alpha p_{an1}, \dots, \alpha p_{anMI})$  до класу  $(KL_l) = (B_{p_{al}})$ , якщо виконується умова –  $(\alpha p_{a1}, \dots, \alpha p_{aMI}) \notin NI_{\mathfrak{R}}$ .

У нашому випадку, отримання досконалої диз'юнктивної нормальної форми (ДДНФ) функції зводиться до побудови ДДНФ для  $F_{KL}$ , яка приймає значення 0 на наборах із  $B_{F_{\overline{KL}}}$  і значення 1 на інших наборах  $E_{KL}^{MI}$ . Після того як досконала диз'юнктивна нормальна форма для  $F_{\overline{KL}}$  отримана, з неї необхідно видалити кон'юнкції  $\mathfrak{R}$ , які не мають властивості  $NI_{\mathfrak{R}} \cap A_{F_{KL}} \neq 0$ .

Наприклад, побудувати ДДНФ логічної функції можна шляхом перетворення кон'юнктивної функції (КФ) виду:

$$D_1 \wedge D_2 \wedge \dots \wedge D_u, \quad (2.6)$$

де  $D_i = p_{ax1}^{\overline{\beta_{i1}}} \vee p_{ax2}^{\overline{\beta_{i2}}} \vee \dots \vee p_{axMI}^{\overline{\beta_{iMI}}}$ ,  $i = 1, 2, \dots, u$  реалізує функцію  $F_{KL}$ ,  $\beta_{iMI}$  – елементи набору  $B_{F_{KL}}$ .

Введемо таке позначення:

$$\overline{p_{ax}^\alpha} = \bigvee_{\beta_i \neq \alpha_i} p_{ax}^\beta. \quad (2.7)$$

Тоді КФ набуває вигляду  $D_1^* \wedge D_2^* \wedge \dots \wedge D_u^*$ ,

де  $D_i^* = \bigvee_{l \neq \beta_{i1}} p_{ax1}^\eta \vee \bigvee_{l \neq \beta_{i2}} p_{ax2}^\eta \vee \dots \vee \bigvee_{l \neq \beta_{iMI}} p_{axMI}^\eta$ ,  $i = 1, 2, \dots, u$ .

Отже, побудова ЕК для класу графічних примітивів, що моделюється, на креслярській документації, можна звести до подальшої реалізації наступних кроків:

- 1) задати характеристичну функцію графічного примітиву, наприклад, для кола це відповідна функція, що описує це коло, для прямокутника або квадрата координати вершин тощо;
- 2) знайти диз'юнктивну нормальну форму ДНФ (або ДДНФ), що реалізує характеристичні функції відповідних примітивів, які знаходяться у проєкції перерізу відповідного об'єкта розпізнавання;
- 3) визначити допустиму (максимальну) кон'юнкцію  $\mathfrak{A}$ , що визначає приналежність об'єкта до класу креслярських примітивів, що розглядається.

Проаналізуємо випадок, коли об'єкти з досліджуваної множини РА описуються показниками, кожен із яких приймає значення з множини  $\{0, 1, \dots, k_{pa} - 1\}$ , тобто, у двійковій формі, зручній для схемо-технічної чи програмної реалізації CNN.

Введемо наступні позначення:

$PA_{mn}^{k_{pa}}$ ,  $k_{pa} \geq 2$  - множина всіх матриць, що характеризують об'єкти аналізу в CNN (існуючі шаблони креслярських примітивів, наприклад,

квадрат-прямокутник або коло-овал, тощо) з елементами  $m \times n$  із  $\{0, 1, \dots, k_{p_a} - 1\}$ ;

$E_k^{r_{p_a}}$ ,  $k_{p_a} \geq 2$ ,  $r_{p_a} \leq n$ , множина всіх  $k_{p_a}$ -их наборів, що мають довжину  $r_{p_a}$ ;

$Q_p(\sigma_{DOP})$ ,  $\sigma_{DOP} \in E_k^{r_{p_a}}$ ,  $\sigma_{DOP} = (\sigma_{DOP_1}, \dots, \sigma_{DOP_r})$ ,  $p \in \{1, 2, \dots, r_{p_a}\}$  - множина всіх наборів виду  $\beta_1, \dots, \beta_r$  в  $E_k^{r_{p_a}}$  яких виконуються умови  $\beta_p \neq \sigma_{DOP_p}$  та  $\beta_j \neq \sigma_{DOP_j}$  при  $j \in \{1, 2, \dots, r_{p_a}\} \setminus \{p\}$ ;

$CU(LU, \sigma_{DOP})$  – множина всіх пар виду  $(HU, \sigma_{DOP})$ ;

$HU - \sigma_{DOP}$  – покриття матриці  $LU \in PA_{mn}^k$ ;

$BU(LU, \sigma_{DOP})$  – множина всіх пар виду  $(HU, \sigma_{DOP})$ ;

$(SU, \sigma_{DOP})$  – сукупність всіх  $\sigma_{DOP}$  – підматриці матриці  $LU$ .

Введемо наступні позначення:

$$CU(LU) = \bigcup_{r_{p_a}=1}^{n=MI} \bigcup_{\sigma_{DOP} \in E_k^r} CU(LU, \sigma_{DOP}), \quad (2.8)$$

$$BU(LU) = \bigcup_{r_{p_a}=1}^{n=MI} \bigcup_{\sigma_{DOP} \in E_k^r} BU(LU, \sigma_{DOP}), \quad (2.9)$$

$$SU(LU) = \bigcup_{r_{p_a}=1}^{n=MI} \bigcup_{\sigma_{DOP} \in E_k^r} \sigma_{DOP}, \quad (2.10)$$

Найбільшу цікавість у задачі синтезу алгоритмів розпізнавання креслярської документації та вирішення задачі подальшої трансформації  $2D \rightarrow 3D$  за допомогою згорткової нейронної мережі представляє дослідження асимптотики значень параметрів  $|CU(LU)|$  та  $|SU(LU)|$ .

Зокрема, необхідно вивчити ситуацію, пов'язану з твердженням, що майже для всіх матриць  $LU$  із  $PA_{mn}^k$  при  $n = MI \rightarrow \infty$  виконується властивість  $\beta$ .

Отже, частка тих матриць із  $PA_{mn}^k$  для яких з  $\varepsilon$  - прийнятною точністю виконується властивість  $\beta$ , прагнути до одиниці (1) і, одночасно,  $\varepsilon \rightarrow 0$  при  $n = MI \rightarrow \infty$ .

Припустимо, що  $PA_{mn}^k = \{LU\}$  простір елементарних подій при розпізнаванні креслярської документації за допомогою CNN.

При цьому кожна подія  $LU$  відбувається з ймовірністю  $1 / |PA_{mn}^k|$ . Математичне очікування випадкової величини  $\hbar(LU)$  позначати як  $M\hbar(LU)$ , а дисперсію –  $D\hbar(LU)$ .

Проаналізуємо, скількома способами можна побудувати матрицю з  $PA = PA_{(nv_1, m\omega_1, \sigma_{DOP_1})} \cap PA_{(nv_2, m\omega_2, \sigma_{DOP_2})}$ .

На першому етапі вибираємо ті елементи, які розташовуються на перетині рядків з номерами з  $nv_1$  та, стовпців з номерами з  $m\omega_1$ . Це завдання можна виконати  $(k_{pa} - 1)^{r_{pa}}$  засобами.

На наступному кроці вибираємо лише ті елементи матриці, які розташовуються на перетині рядків з номерами з  $nv_2$ , та стовпців з номерами з  $m\omega_2$ .

Необхідно врахувати, що  $ab$  з них розташовані одночасно на перетині рядків з номерами із  $nv_1$  та стовпців з номерами з  $m\omega_2$ ,  $((k_{pa} - 1)^{l-a} - \text{способів})$ .

У загальному випадку, довільно довизначивши рядки матриці з номерами з  $nv_1 \cup nv_2$   $(k_{pa}^{(r_{pa}+l-a)MI+ab-r_{pa}^2-l^2} - \text{способів})$ , вибираємо рядки, що залишилися  $(k_{pa}^{PA \cdot MI - (r_{pa}+l-a)MI} - \text{способів})$ .

Припустимо, що для всіх матриць, що характеризують об'єкти аналізу в CNN (показників (ознак графічних примітивів, що знаходяться на креслярській документації, обраної для розпізнавання)  $LU \in PA_{mn}^k$  при  $n = MI \rightarrow \infty$  справедливо відношення:

$$|SU_1(LU)| = 0.$$

Коли на підставі лем, розглянутих у роботах [77-81], проаналізуємо такі ситуації.



1. Якщо  $n^\alpha \leq m \leq k^{n^\beta}$ ,  $\alpha > 1$ ,  $\beta < 1$ , то має місце  $PA\eta_1(LU) \approx$

$$PA\eta_2(LU) \approx \sum_{r_{pa} \in \lambda_1} CU_m^{r_{pa}} \cdot CU_n^{r_{pa}} r_{pa} (k_{pa} - 1)^{r_{pa}} \cdot k_{pa}^{r_{pa} - r_{pa}^2}$$

при  $n = MI \rightarrow \infty$ ;

2. Якщо  $n^\alpha \leq m \leq k^{n^\beta}$ ,  $\alpha > 1$ ,  $\beta < 1$ , то має місце

$$\frac{D\eta_2(LU)}{PA(\eta_2(LU))^2} \rightarrow 0 \text{ при } n = MI \rightarrow \infty,$$

де  $\eta_{(nv, m\omega)}(LU, \sigma_{DOP})$  – випадкова величина дорівнює 1, якщо  $LU \in$

$PA_{(nv, m\omega, \sigma_{DOP})}$  і 0 в іншому випадку, маємо

$$PA(\eta_3(LU)) = \sum_{r_{pa} \geq r_{pa1}} \sum_{\substack{nv \in V_r^m \\ m\omega \in W_r^n}} \sum_{\sigma_{DOP} \in E_k^r} P(\eta_{(nv, m\omega)}(LU, \sigma_{DOP}) = 1), \quad (2.11)$$

де  $nv \in V_{r_{pa}}^m$ ,  $m\omega \in W_{r_{pa}}^n$ ,  $\sigma_{DOP} \in E_k^{r_{pa}}$ ;

$P(\eta_{(nv, m\omega)}(LU, \sigma_{DOP})) = 1$  – ймовірність того, що

$\eta_{(nv, m\omega)}(LU, \sigma_{DOP}) = 1$ .

Отже, в силу,

$$|PA_{nv, m\omega, \sigma_{DOP}}| = (k_{pa} - 1)^{r_{pa}} \cdot k_{pa}^{mn - r_{pa}^2},$$

отримаємо

$$PA\eta_3(LU) = \sum_{r_{pa} \geq r_{pa1}} CU_{n=MI}^{r_{pa}} \cdot CU_m^{r_{pa}} \cdot r_{pa}! (k_{pa} - 1)^{r_{pa}} \cdot k_{pa}^{r_{pa} - r_{pa}^2}, \quad (2.12)$$

Так як при  $r_{pa} \geq r_{pa1}$

$$CU_{n=MI}^{r_{pa}} \cdot CU_m^{r_{pa}} \cdot r_{pa}! (k_{pa} - 1)^{r_{pa}} \cdot k_{pa}^{r_{pa} - r_{pa}^2} \leq \frac{(mn)^{r_{pa}}}{r_{pa}!} \cdot r_{pa}^{2 \cdot r_{pa} - r_{pa}^2} \leq \left(\frac{k_{pa}^2 \cdot e}{r_{pa}}\right)^{r_{pa}}, \quad (2.13)$$

то за досить великої кількості об'єктів (графічних примітивів) описаних за допомогою своїх «матриць показників» і  $n = MI \rightarrow \infty$  матимемо

$$\sum_{r_{pa} \geq r_{pa1}}^{n=MI} CU_{n=MI}^{r_{pa}} \cdot CU_m^{r_{pa}} \cdot r_{pa}! (k_{pa} - 1)^{r_{pa}} \cdot k_{pa}^{r_{pa} - r_{pa}^2} \leq n \left( \frac{k_{pa}^2 \cdot e}{\log_{k_{pa}} nm} \right)^{\log_{k_{pa}} nm} \rightarrow 0. \quad (2.14)$$

Таким чином,

$$|BU(LU)| \leq CU_{n=MI}^{\lceil \log_{k_{pa}} nm \rceil} mn \leq \frac{n^{\log_{k_{pa}} nm} \cdot mn}{\lceil \log_{k_{pa}} nm \rceil!}. \quad (2.15)$$

Також маємо:

$$|SU(LU)| \geq \sum_{r_{pa} \in \lambda_1} \frac{(nm)^{r_{pa}-1}}{r_{pa}!} \left(1 - \frac{r_{pa}}{n}\right)^{r_{pa}} \cdot \left(1 - \frac{r_{pa}}{m}\right)^{r_{pa}} \geq \sum_{r_{pa} \in \lambda_1} \frac{(nm)^{r_{pa}-1}}{\lceil \log_{k_{pa}} nm \rceil!}, \quad (2.16)$$

$$\text{де } r_{pa1} = \frac{1}{2} \cdot \log_{k_{pa}} nm - \frac{1}{2} \cdot \log_{k_{pa}} \log_{k_{pa}} nm - \log_{k_{pa}} \log_{k_{pa}} \log_{k_{pa}} n;$$

$\lambda_1$  – інтервал.

Звідси маємо

$$\frac{|SU(LU)|}{|BU(LU)|} \geq \frac{(mn)^{r_{pa1}-2}}{n^{\log_{k_{pa}} nm}} \geq n^{(\alpha+1)(r_{pa1}-2) - \log_{k_{pa}} nm} \rightarrow \infty \quad (2.17)$$

при  $n = MI \rightarrow \infty$ .

Відповідно, коли кількість рядків матриці, що містить типові показники об'єктів (або креслярських примітивів), по порядку наявна більша кількість стовпців, величина  $|SU(LU)|$  по порядку більша за величину  $|BU(LU)|$ .

Креслярські графічні примітиви, які покладені в основу застосування модуля ЕК, можна розділити, застосувавши критерій їх приналежності до джерела отримання інформації, наприклад, зовнішні та внутрішні. У цьому випадку кожен клас креслярських графічних примітивів, які використовуються в ЕК, розіб'ємо на групи, і підгрупи за параметрами (ознаками), див. рис. 2.5.

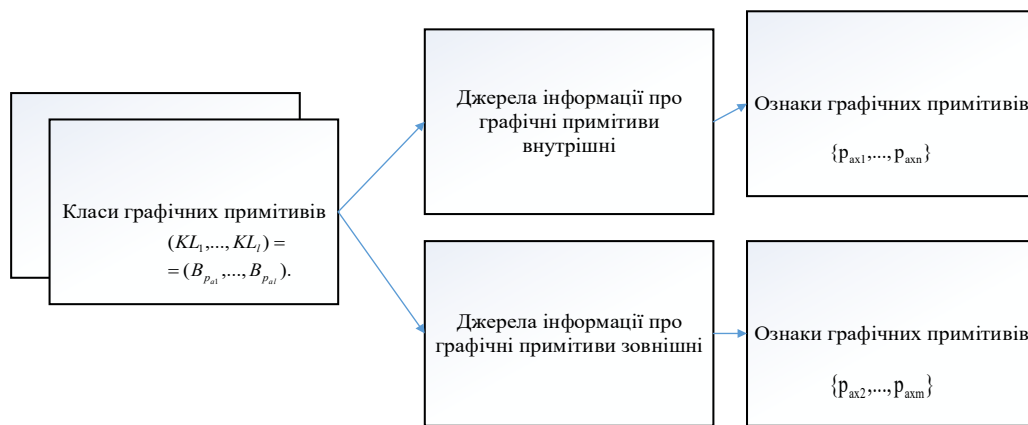


Рисунок 2.5 Модель класифікації «Джерела інформації про графічні примітиви».

*Джерело: побудовано автором*

Слід пам'ятати, що всі графічні примітиви додатково характеризуються атрибутивною та метричною частинами. Атрибутивна частина параметрів графічного примітиву характеризує форми об'єктів, наприклад, лінії та стрілки; прямокутники; кола, еліпси, дуги, сегменти та сектори; криві; сполучні лінії; тривимірні об'єкти (куби, кулі, циліндри тощо); текстові дані. Метрична дає можливість зіставляти розміри різних частин об'єкта один одному, і навіть розмірам інших об'єктів на кресленні. При цьому частина даних може знаходитися у зовнішніх джерелах, наприклад, частина текстової інформації, що описує об'єкт, може розташовуватися на інших джерелах, а не безпосередньо на кресленні, що розпізнається.

Автоматизація пошуку нових типів зображень, що розпізнаються на технічно-креслярській документації, передбачає застосування моделей, що базуються на комбінаторному аналізі описів ознак  $\{sp_{a1}, \dots, sp_{am}\}$ .

У роботах багатьох учених [73-81] було показано, що ці методи дають прийнятний результат розпізнавання в ситуації, коли дані представлені в дискретній формі і кількість значень ознаки, що допускаються, невелика [78] (див. рис. 2.6).

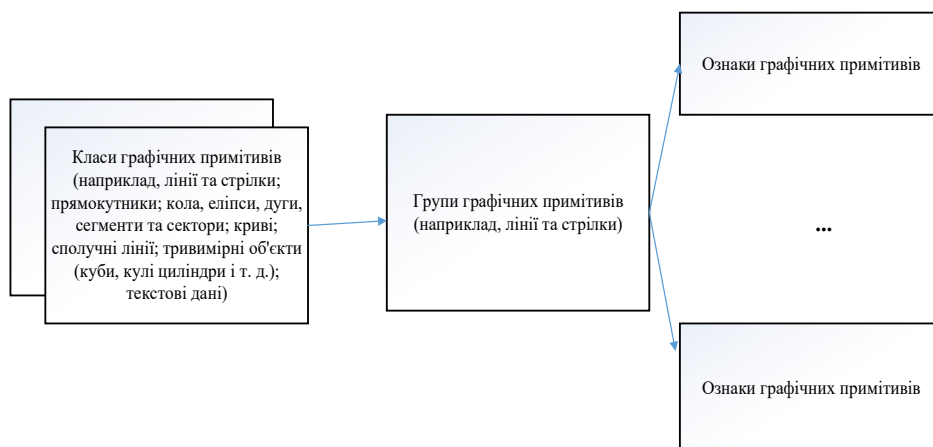
Першим етапом будь-якого процесу розпізнавання креслярсько-технічної документації є розкладання вихідного графічного об'єкта,

наприклад, креслення об'єкта на примітиви. На цьому етапі неминуче виникають складнощі, пов'язані з визначенням місць «розрізу» графічних образів при різноманітті варіантів. І це породжує неминучість різноманіття варіантів залучення примітивів. Наприклад, при наявних видах графічного об'єкта, показаних на рисунку 2.7, і можливих неточностях при розпізнаванні деталей, можливі різні трактування конструкції, її трансформації  $2D \rightarrow 3D$ . В результаті можуть бути отримані різні варіанти при 3D моделюванні.

Друга складність – визначити необхідну мінімальну множину графічних примітивів. У сучасних програмних продуктах, призначених для роботи з креслярською документацією, є характерна особливість, яка полягає в необхідності подати аналізований об'єкт як набір числових даних.

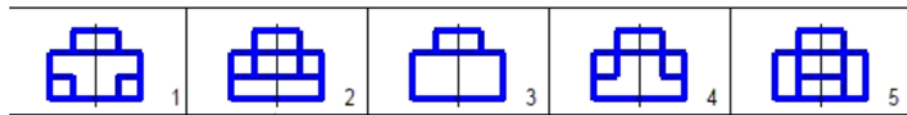
Такими поняттями як: лінії та стрілки; прямокутники; кола, еліпси, дуги, сегменти та сектори; криві; сполучні лінії; тривимірні об'єкти (куби, кулі циліндри тощо); текстові дані ні комп'ютери ні прикладне програмне забезпечення не оперує.

Тому в роботах [73-78] і робиться акцент на застосування апарату цифрових ЕК для роботи з графічними примітивами об'єктів.

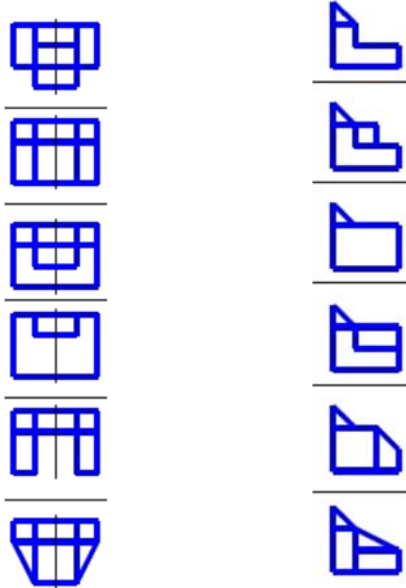


**Рисунок 2.6. Структура класифікації графічних примітивів для проєктування ЕК.**

*Джерело: побудовано автором*



А) Головний вигляд



Б) Вид зверху

В) Вид зліва

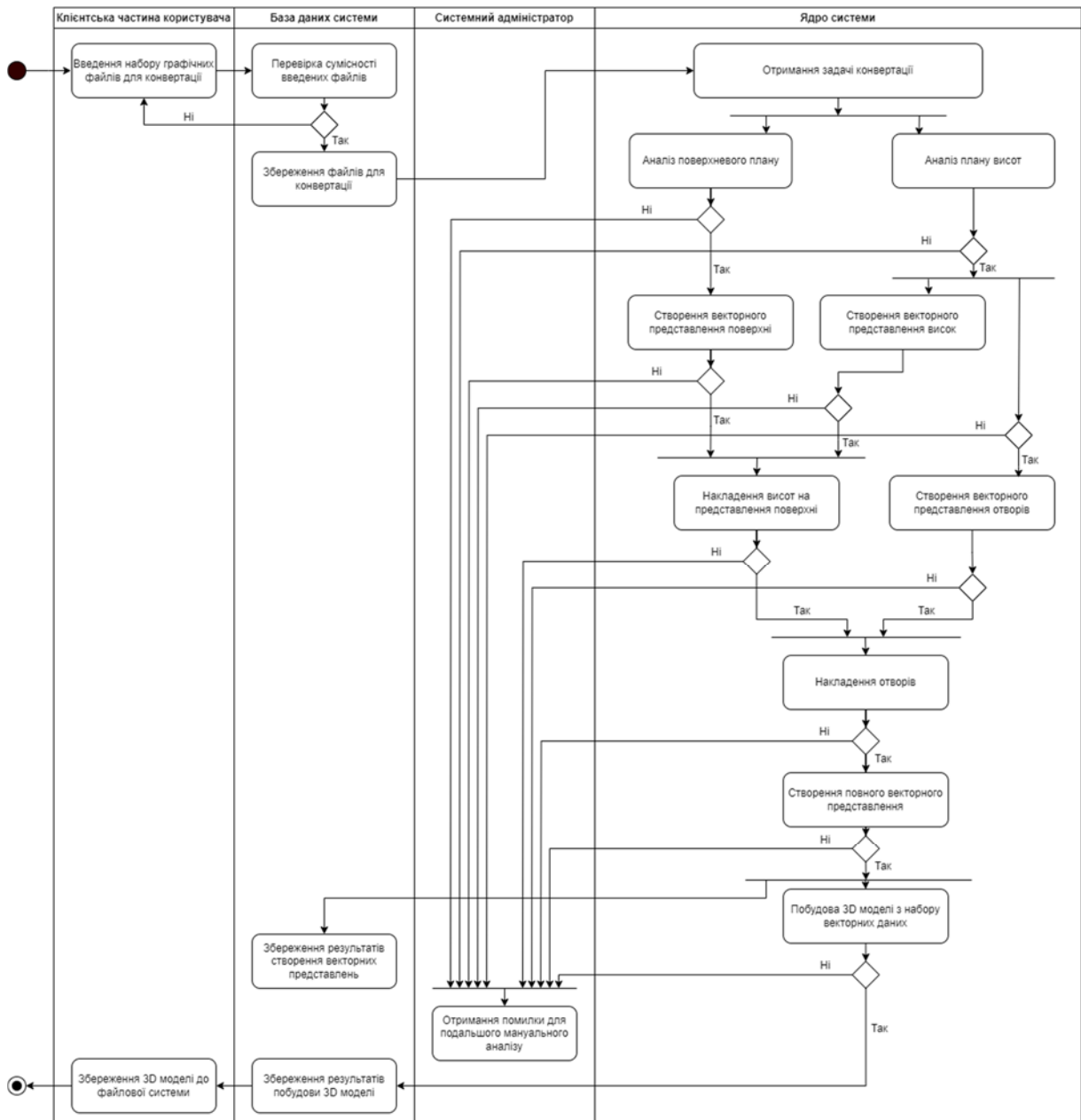
**Рисунок 2.7. Наявні види графічного об'єкта, що складаються з ЕК у вигляді прямокутників.**

*Джерело: побудовано автором*

Вважатимемо, що класифікація об'єктів (тобто графічних примітивів, що використовуються на креслярсько-будівельній документації), представляє собою множини потенційних варіантів зображень проєкцій, властивих тій чи іншій номенклатурі графічних об'єктів.

У запропонованій CNN (рис.2.8) кожен елементарний класифікатор буде відносити кожен умовний квадрат до певного класу креслярських примітивів. На стадії навчання CNN необхідно максимізувати кількість варіантів регіонів-кандидатів, задіяних у процесі розпізнавання примітивів на креслярсько-технічній документації. Далі за допомогою модуля з ЕК виконується операція зі складання карти об'єктів. Причому дана карта складається з урахуванням всіх пропозицій щодо наявних у базі примітивів. Оскільки не ставиться технічне завдання програмної реалізації CNN можна

на даному етапі досліджень скористатися однією з відкритих бібліотек мови програмування Python, наприклад, PyTorch [82, 83].



**Рисунок 2.8. Алгоритм функціонування нейронної мережі інформаційної технології розпізнавання графічної інформації**

*Джерело: побудовано автором*

Більшість відкритих бібліотек, що використовуються для навчання CNN, написані мовою Python або Go, тому імплементація в таку мережу модуля, що реалізує рішення задачі, класифікації креслярських примітивів,

також може бути виконана на Python або Go. Нижче наведено фрагмент такого ЕК, створеного на основі застосування бібліотеки TensorFlow [84, 85], див. рис. 2.9.

Показаний фрагмент ЕК, призначений для вирішення наступних завдань:

- 1) Виконати аналіз зображення елемента креслення та знайти на ньому елементарний примітив, наприклад, коло, яке відповідає отвору, як це показано на рисунку 2.4;
- 2) якщо отвір виявлено, необхідно визначити його категорію. Наприклад, це може бути наскрізний отвір деталі або поглиблення;
- 3) Розрахувати ступінь впевненості у зробленому висновку.

```
1 import tensorflow as tf
2 import sys
3 from numpy import argsort
4
5 image_path = sys.argv[1]
6
7 # зчитуємо файл image_data
8 image_data = tf.gfile.FastGFile(image_path, 'rb').read()
9
10 # завантажує вибраний файл та видаляє символ розрива строки
11 label_lines = [line.rstrip() for line in tf.gfile.GFile("/tf_files
    /retrained_labels.txt")]
12
13 # відокремлює граф від файла
14 with tf.gfile.FastGFile("/tf_files/retrained_graph.pb", 'rb') as f:
15     graph_def = tf.GraphDef()
16     graph_def.ParseFromString(f.read())
17     _ = tf.import_graph_def(graph_def, name='')
18
19 # завантажує image_data як
20 with tf.Session() as sess:
21     softmax_tensor = sess.graph.get_tensor_by_name('final_result:0')
22     predictions = sess.run(softmax_tensor, {'DecodeJpeg/contents:0':
        image_data})
```

**Рисунок 2.9. Фрагмент коду елементарного класифікатора креслярських примітивів на основі застосування бібліотеки TensorFlow.**

*Джерело: побудовано автором на основі джерел [84, 85]*

Переваги логічних процедур навчання при конструюванні елементарних класифікаторів, що використовуються в задачі розпізнавання креслень та трансформації  $2D \rightarrow 3D$ , є:

- можливості отримувати функції класифікації креслярських примітивів у рамках залучення CNN з модулем ЕК графічних примітивів для розпізнавання креслярсько-технічної документації з мінімальним рівнем помилки процедури класифікації;
- роботою зі складно структурованими даними за рахунок використання ЕК, що описують, наприклад, штрихування на проєкції графічного об'єкта;
- швидка заміна правил  $DR(p_{axi})$  у разі зміни структури даних, що аналізуються, без модифікації самого алгоритму розпізнавання креслярсько-технічної документації.

З урахуванням проаналізованих у першому розділі дисертації робіт, і беручи до уваги основну концепцію обробки та аналізу зображень, яка спирається на фундаментальну роботу [86], можна подати етапи процедури розпізнавання креслень у вигляді трирівневої моделі, як це показано на рисунку 2.10.

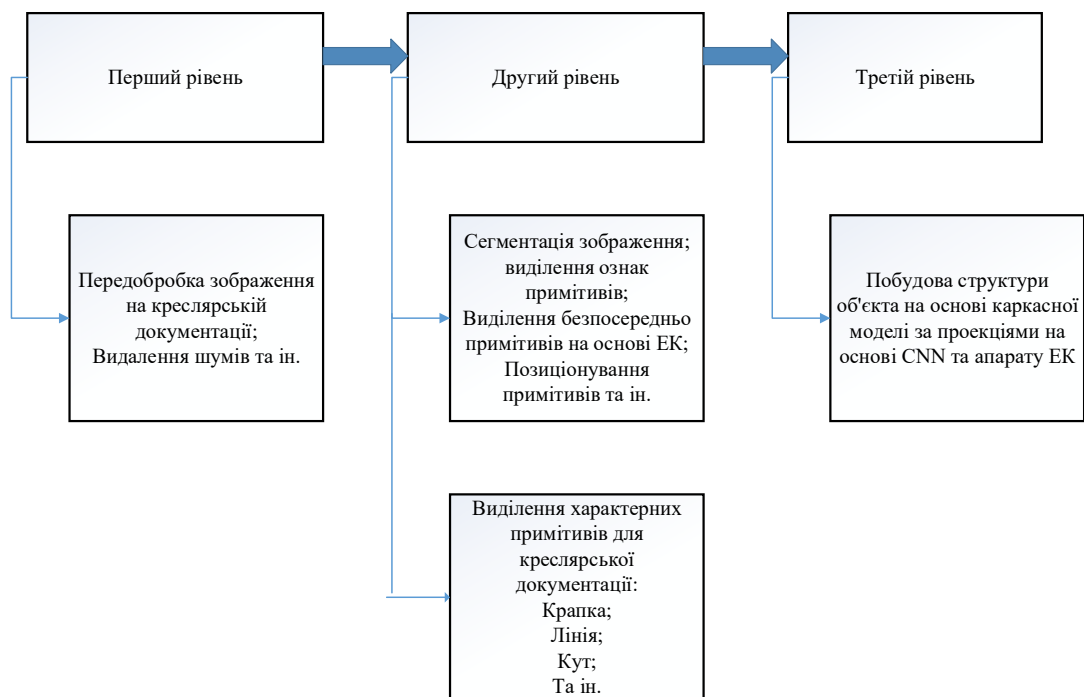


Рисунок 2.10. Трирівнева модель обробки креслень на основі застосування апарату ЕК та CNN.

Джерело: побудовано автором



## 2.2. Алгоритм трансформації $2D \rightarrow 3D$ на основі інформації з розпізнаного креслення

Проекції графічного об'єкта, які представлені на кресленні, несуть найбільш важливу інформацію для отримання 3D образу графічного об'єкта. Однак такої інформації достатньо для простих графічних об'єктів, проєкції яких досить добре описуються наявними елементарними класифікаторами та примітивами, наприклад, для проєкцій, показаних на рис. 2.11.

Як показав огляд попередніх досліджень, див. розділ 1, досі не створено ефективного алгоритму, що дозволяє вважати і проаналізувати різноваріантну графічну інформацію, яка може траплятися на креслярсько-технічній документації.

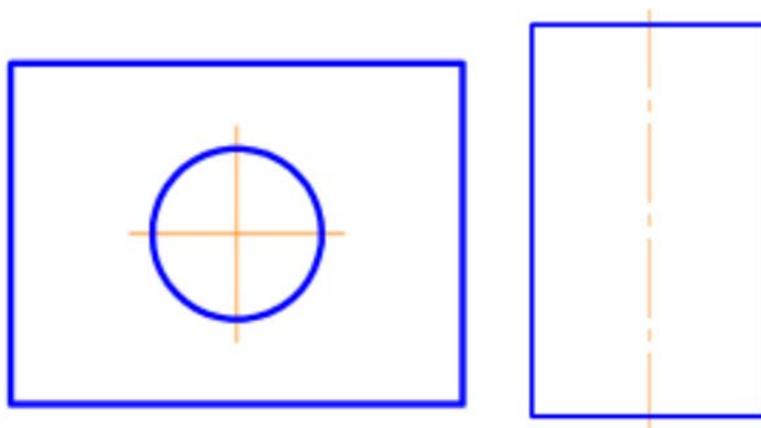
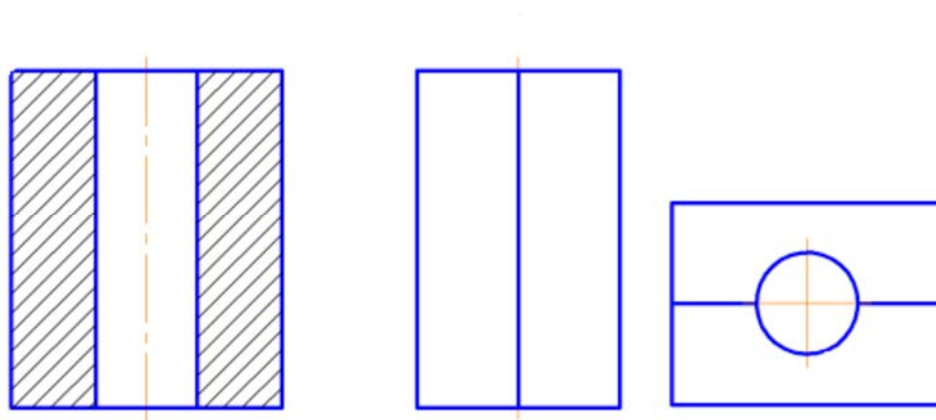


Рисунок 2.11. Приклад простих проєкцій графічного об'єкта, необхідних для трансформації  $2D \rightarrow 3D$ .

*Джерело: побудовано автором*

Однак, у більшості ситуацій проєкції графічних об'єктів містять додаткову інформацію, наприклад, штрихування тощо (див. рис. 2.12).



**Рисунок 2.12. Приклад простих проєкцій графічного об'єкта з додатковою інформацією (штрихуванням).**

*Джерело: побудовано автором*

Як показали результати виконаного аналізу попередніх досліджень, робота швидше ведеться у напрямі, який дозволить покроково розширювати класи графічної інформації, яка використовується та доступна для обробки. Зокрема, рішенню саме цього завдання і присвячено попередній підрозділ роботи.

Як було показано в роботах [20-22] після отримання попередньої каркасної моделі графічного об'єкта на підставі цифрового креслення або розпізнаних проєкцій графічного об'єкта на паперовому носії та видалення з неї хибних елементів, можна говорити про якісну трансформацію  $2D \rightarrow 3D$ .

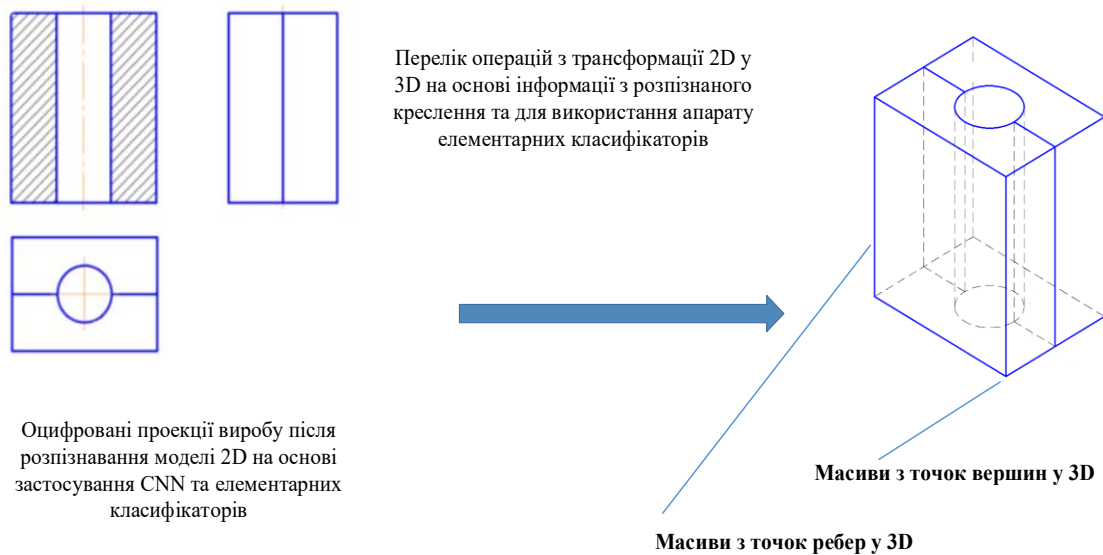
В рамках цього підрозділу дисертації стоїть завдання розвитку існуючого алгоритму зчитування проєкцій з креслення з урахуванням застосування процедури залучення ЕК для ситуацій, коли на проєкціях є розрізи. Відповідно є перерізи, що дещо ускладнює процес трансформації  $2D \rightarrow 3D$  на основі каркасної моделі графічного об'єкта.

В даному випадку ознаки, які можна використовувати при побудові елементарних класифікаторів, можуть бути такими. Наявність на проєкції/проєкціях графічного об'єкта креслення:

- 1) штрихування;
- 2) ліній обриву;
- 3) позначення «повернуто»;
- 4) розриву між частинами проєкції;
- 5) інші показники та креслярські елементи.

Враховуючи вищесказане, можна модифікувати алгоритм [20-22] трансформації  $2D \rightarrow 3D$  на основі інформації з розпізнаного креслення, враховуючи можливість застосування апарату ЕК, наприклад для вирішення таких підзадач, як: перевірка наявності на проєкціях штрихування, ліній обриву, позначення «повернуто», розриву між частинами проєкції тощо (див. рис. 2.13)

Модифікований алгоритм трансформації  $2D \rightarrow 3D$  на основі інформації з розпізнаного креслення та для використання апарату ЕК показаний на рисунку 2.14.



**Рисунок 2.13. – Спрощена схема реалізації процедури трансформації  $2D \rightarrow 3D$  на основі інформації з розпізнаного креслення та для використання апарату ЕК**

*Джерело: побудовано автором*

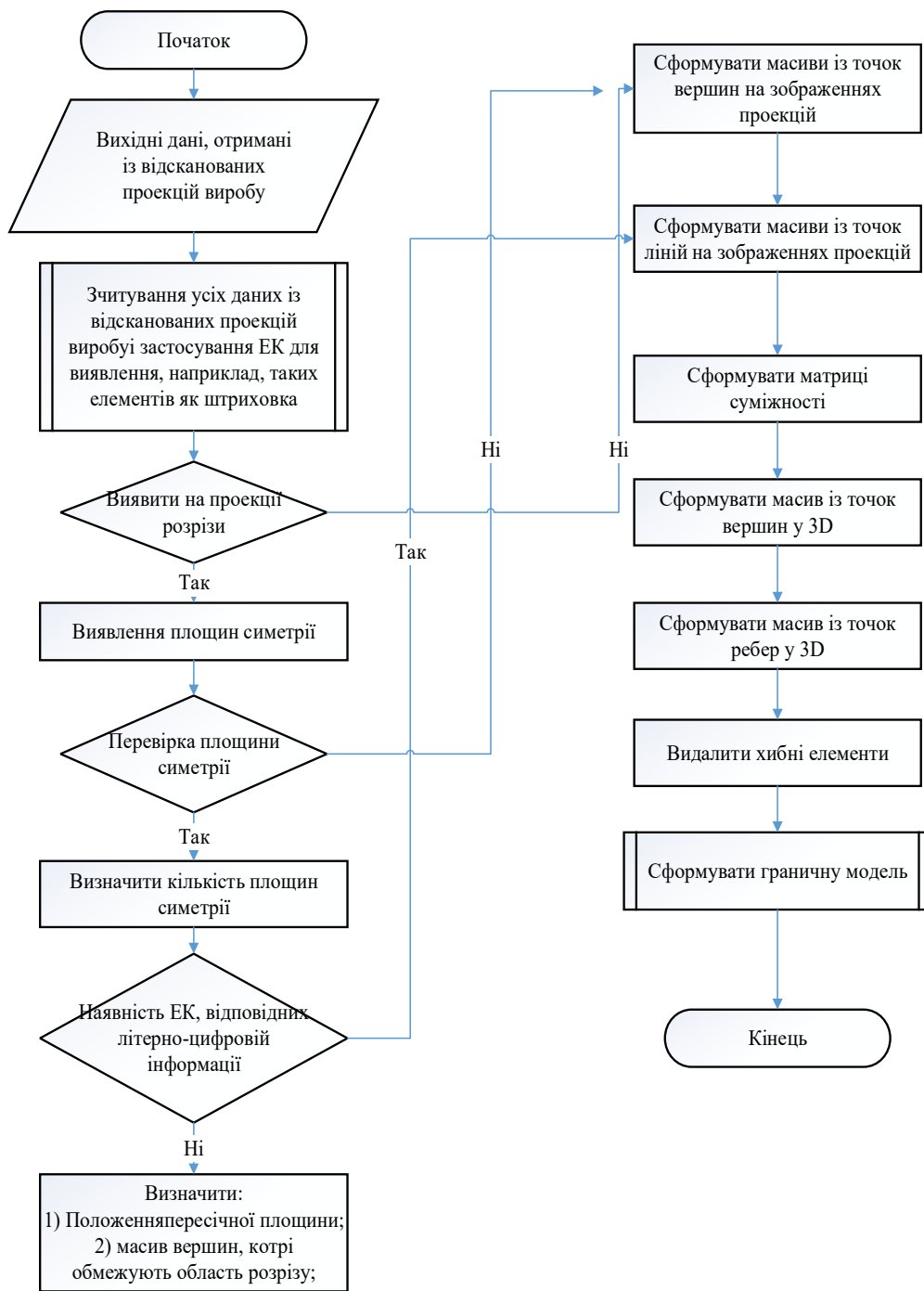
Для простоти демонстрації роботи модифікованого алгоритму трансформації, об'єкт, показаний у лівій частині рисунка 2.7, підібраний досить простим. Такий об'єкт повністю характеризується трьома наявними проєкціями, показаними рисунку 2.7. На одній із проєкцій є штрихування, відповідна до ЕК – Штрихування. ЕК штрихування та розріз містять інформацію про відсутність матеріалу, що важливо для реалізації трансформації  $2D \rightarrow 3D$ . Ця інформація необхідна для видалення помилкових елементів із моделі 3D.

При роботі алгоритму, блок схема якого показана на рисунку 2.10, сформуємо три матриці суміжності. Кожна матриця відповідає своїй проєкції графічного об'єкта. Далі необхідно, наприклад, за допомогою згорткової нейронної мережі, описаної в попередньому параграфі роботи, сформувати відповідні масиви точок. Ці масиви опишуть вершини графічного об'єкта та його ребра. Далі на основі цієї інформації можна сформувати масив для точок, які відповідають положенню вершин та ребер графічного об'єкта у просторі 3D.

Видалення помилкових елементів відбувається аналогічно до процедури, описаної в класичних роботах [20-22]. На підставі даних з ЕК, що містять інформацію про наявні розрізи та порожнини у графічного об'єкта, отримаємо каркасну модель, яка власне і завершує процедуру трансформації  $2D \rightarrow 3D$ .

Таким чином, розширивши вхідну інформацію за рахунок імплементації в базовий алгоритм [20-22] апарату ЕК, це дозволяє ефективно виконувати трансформацію  $2D \rightarrow 3D$ , спрощуючи процес видалення помилкових об'єктів.

На погляд автора, задіяння апарату ЕК в алгоритмі трансформації  $2D \rightarrow 3D$  на основі інформації з розпізнаного креслення, на якому є прості розрізи, дасть можливість виконувати трансформацію  $2D \rightarrow 3D$  з меншим числом хибних елементів, а також знизити обчислювальні ресурси на виконання даної трансформації.



**Рисунок 2.14. Блок-схема алгоритму трансформації 2D → 3D на основі інформації з розпізнаного креслення та для використання апарату ЕК.**

*Джерело: побудовано автором на основі джерел [20-22]*

## ВИСНОВКИ ДО РОЗДІЛУ 2

У результаті проведених у цьому розділі досліджень, зроблено такі висновки:

- Запропоновано архітектуру згорткової нейронної мережі з модулем ЕК графічних примітивів для вирішення задачі розпізнавання креслень та трансформації  $2D \rightarrow 3D$ .
- Розроблено модель класифікатора графічних зображень на основі покриттів класів та елементарних класифікаторів примітивів для підвищення ефективності навчання CNN.
- Показано, що парадигмою побудови бази примітивів на навчання CNN (чи логічних процедур навчання - ЛПН) є пошук інформативних фрагментів описів відповідних графічних примітивів. Ці фрагменти при створенні конкретних проектних рішень для систем розпізнавання креслярсько-будівельної документації, наприклад, на основі CNN, дозволять однозначно робити висновок про наявність (або відсутність) в базі системи розпізнавання певного графічного примітиву в рамках класу.
- Викладено основні принципи конструювання ЛПН з використанням апарату логічних функцій, що дозволить на практиці створювати ефективні програмні рішення CNN у задачах розпізнавання креслень та трансформації  $2D \rightarrow 3D$ ;
- Показано, що конструювання множини ЕК для креслярських примітивів зводиться до знаходження допустимих та максимальних кон'юнкцій для функції, що описує аналізований клас графічних об'єктів.

Основні результати розділу висвітлені у науково-дослідній роботі НДР № 0122U001549 «Моделювання інтелектуальних систем управління діяльністю підприємств», (довідка від 29.08.2023 №1453/24)

Основні результати розділу опубліковані в наукових працях автора:

Наукові праці, в яких опубліковані основні наукові результати дисертації:

[1], [5], [6], [7], [9].

Наукові праці, які засвідчують апробацію матеріалів дисертації:

[12], [13].

Наукові праці, які додатково відображають наукові результати дисертації

[16], [17], [18], [19], [20].

## **РОЗДІЛ 3.**

# **МОДЕЛЮВАННЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ РОЗПІЗНАВАННЯ ГРАФІЧНОЇ ІНФОРМАЦІЇ ДЛЯ СТВОРЕННЯ 3D МОДЕЛІ НА ОСНОВІ 2D ЗОБРАЖЕННЯ**

### **3.1. Концептуальна модель інформаційної технології розпізнавання графічної інформації**

Розробка концептуальної моделі надає чітке розуміння проблемної області та викликів, пов'язаних із перетворенням 2D зображень у 3D модель на основі нейронної мережі та визначити вимоги до вихідних даних і взаємодію між різними складовими моделі.

Концептуальна модель забезпечує високорівневе представлення архітектури системи та її компонентів, також розробку загальної структури системи, включаючи модель нейронної мережі, потік даних і взаємодію між різними складовими.

Інформаційна технологія розпізнавання графічної інформації для створення 3D моделі на основі 2D зображення може включати такі складові (рис.3.1.):

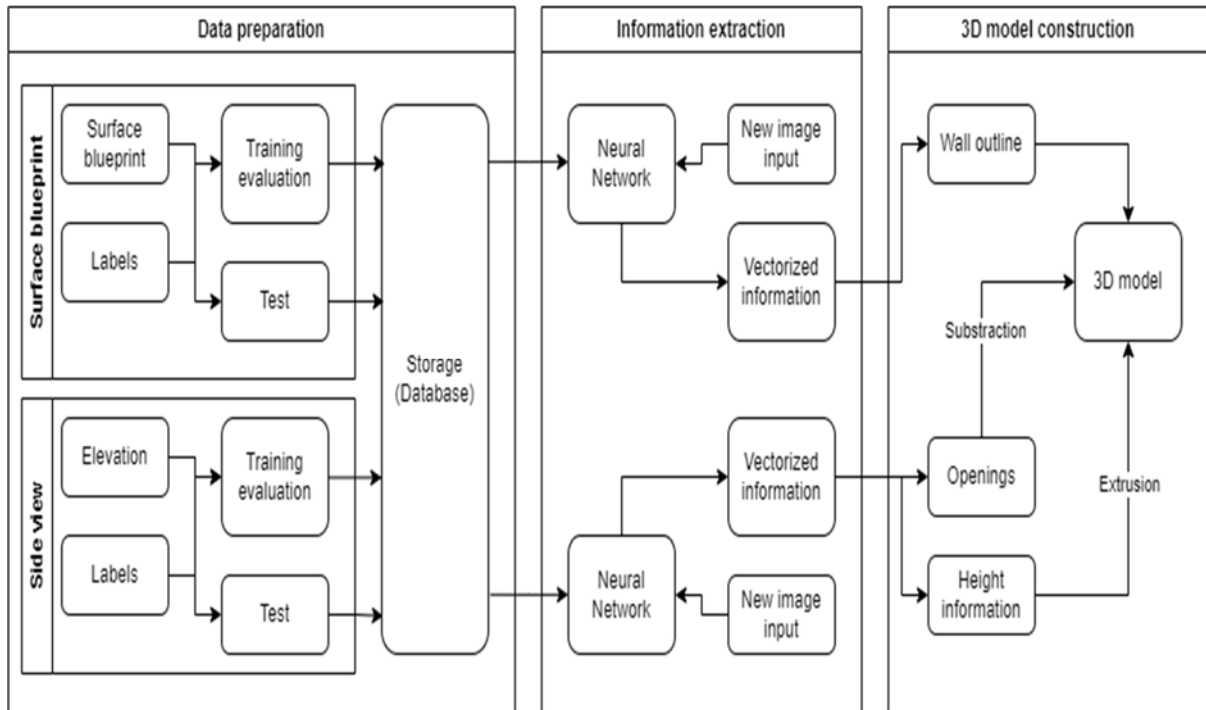
1. Data preparation (Підготовка даних)
2. Information extraction (Обробка інформації)
3. 3D model construction (Побудова 3D моделі)

Важливо відмітити, що одним із ключових моментів для впровадження згорткової нейронної мережі з модулем елементарних класифікаторів графічних примітивів у завданнях розпізнання графічної інформації (креслярської документації) є процес підготовки даних. Особливістю такої підготовки даних є специфічність необхідних даних. На відміну від існуючого класичного розпізнавання достатньо було б стандартно промаркованих даних, а в нашому випадку вперше постає необхідність отримання додаткових даних:

- набір даних плану площини (2D креслення);



- набір даних про висоту;
- доповнення даних та виділення креслярських елементів (лейблінг).



**Рисунок 3.1. Концептуальна модель інформаційної технології розпізнавання графічної інформації на основі нейронної мережі**

*Джерело: побудовано автором*

Запропонований інноваційний підхід, що спроектований у вигляді вище зазначеної концептуальної моделі, зображеної на рис. 3.1, включає обробку та аналіз графічної інформації за допомогою глибокого навчання, що відрізняється від традиційного монокулярного відновлення глибини.

Розглянемо дану концептуальну модель відносно етапів (по вертикалі) та процесів (по горизонталі).

Етапи концептуальної моделі інформаційної технології розпізнавання графічної інформації на основі нейронної мережі:

*Data preparation (Підготовка даних).* У процесах обробки даних використовуються ідентичні паралельні процедури, які включають

мануальний процес лейблінгу (якщо це необхідно), тренування та тестування нейронних мереж.

*Information extraction (Отримання даних).* Використовуються дві окремі нейронні мережі для обробки різних видів зображень: плану поверхні та плану висот. Подібна архітектура нейронної мережі, оптимізована для розпізнавання специфічних характеристик зображення, що надає додаткову гнучкість для вибору зображення для аналізу.

*3D model construction (Побудова 3D моделі).* Отримані векторизовані дані об'єднуються та використовуються для створення 3D моделі. Цей процес включає семантичну сегментацію та інтерпретацію даних для побудови каркасу моделі, а також застосування додаткових даних про отвори для отримання більш деталізованого представлення.

Розглянемо кожен процес розпізнавання та створення 3D моделей на основі графічної інформації за допомогою нейронних мереж:

*Підготовка даних (Surface blueprint, Side view).* Процеси Surface blueprint та Side view ідуть паралельно і відрізняються лише вхідними даними про графічну інформацію. Графічна інформація, яка є вхідною в процесі Surface blueprint – 2D креслення, а в процесі Side view – висоти. На цьому етапі зображення підлягають процесам очищення та попередньої обробки, включаючи лейблінг, який зазвичай виконується вручну. Потім, дана графічна інформація використовується для тренування нейронної мережі. Звичайно, цей процес також включає розділення набору даних на тренувальний та тестовий набори для валідації моделі. Результати та стан моделі зберігаються в Storage (Database) для подальшого використання.

*Отримання інформації.* На даному етапі можна розглянути концептуальну модель з точки зору автоматизованої системи та користувачької. З точки зору системи: нейронна мережа аналізує графічну інформацію та екстракує корисну інформацію, перетворюючи її в векторизовану форму. Цей процес включає в себе побудову глибоких ознак графічних об'єктів і використання їх для створення деталізованих векторів.

З точки зору користувача на даному етапі графічна інформація про об'єкт може бути замінена без повторного тренування нейронної мережі, якщо нове зображення є варіацією оригінального зображення.

*Генерування 3D моделі.* На цьому етапі векторизована інформація з плану висот розділяється на два набори даних: інформація про отвори та інформація про висоти. З плану площини виокремлюється інформація про контури поверхонь. Далі, за допомогою цієї інформації створюється каркас 3D моделі. Після отримання проміжного результату застосовується інформація про отвори, що призводить до отримання фінальної векторизованої інформації для побудови 3D моделі.

Дана модель відрізняється від існуючих подібних аналогів саме отриманою інформацією, яка буде більш повною і більш широко розкриватиме можливість перенести точні деталі для 3D моделі.

Відповідно кожного етапу концептуальної моделі визначимо методики:

1) *Data preparation* – методика підготовки даних для навчання нейронної мережі.

2) *Information extraction* – методика навчання та тестування нейронної мережі.

3) *3D model construction* – концепція побудови 3D моделі.

Також для кожного етапу визначимо методи:

1) *Data preparation* – методи машинного навчання, методи нейронних мереж, включаючи глибокі згорткові нейронні мережі (DCNN) та багат шарові перцептрони (MLP), метод послідовних ітерацій навчання на основі попередніх результатів.

2) *Information extraction* – методи комп'ютерного зору та моделювання, методи генерації векторизованих представлень

3) *3D model construction* – методи алгоритмів комп'ютерного зору та машинного навчання.

Опишемо принцип використання вище зазначених методів та методик інформаційної технології розпізнавання графічної інформації на основі нейронної мережі.

*Методики інформаційної технології розпізнавання графічної інформації на основі нейронної мережі* (пояснимо на прикладі архітектурних об'єктів будівельної галузі):

1) *Data preparation* – Методика підготовки даних для навчання нейронної мережі.

Існують різні типи архітектурних креслень, таких як плани поверхів, розрізи та фасади, які є семантично насиченими та їх достатньо для перетворення на 3D-модель. Кожне з них можна легко отримати з 3D-моделі, тоді як важко, але не неможливо відновити модель за допомогою цих креслень. Використання лише планів поверхів не задовольняє повну реконструкцію 3D-моделі будівлі, оскільки з них не можна інтерпретувати інформацію про висоту та розташування отворів. Додаткове джерело даних слід передбачити та використовувати для зручного й автоматизованого створення моделі.

Серед інших типів креслення фасадів відображають виключно те, що є на планах поверхів, оскільки вони систематично посиляються один на одного. З іншого боку, розрізи та перспективи складно зв'язати безпосередньо з планами поверхів навіть для кваліфікованого спеціаліста. План поверху не обов'язково передбачає деталі секцій. Подібним чином можуть бути відсутні всі компоненти плану поверху. Таким чином, плани поверхів і креслення висоти перетворюються з потенційних джерел даних у набори даних для відтворення таких моделей. Реконструкція 3D-моделі без будь-якого втручання людини вимагає автоматизації кожного кроку, тому очевидно, що для автоматичного розпізнавання компонентів має бути набір даних для навчання нейронної мережі. Незважаючи на те, що наборів даних растрових планів поверхів небагато, немає ані наборів даних

растрового креслення висот, ані міток для таких зображень. Тому план поверху та набори даних висоти з відповідними мітками готуються з нуля для цілей дослідження. Зображення двовимірних креслень архітектурних будинків беруться з наявних наборів даних або онлайн-платформ, але всі мітки проставляються вручну. Далі набори даних поділяються на навчальні, перевірочні та тестові. Навчальний набір даних використовується для вивчення особливостей вибірок даних, що впливає з назви. Набір перевірки застосовується для перевірки того, чи алгоритм дає збій під час вивчення особливостей. Нарешті, набір тестів використовується для перевірки, чи алгоритм вивчає особливості належним чином.

*План поверху* — це креслення в масштабі, що показує просторове розташування. Деталі плану поверху змінюються відповідно до потреб проекту. Наприклад, концептуальний план поверху може включати лише контур стіни, тоді як креслення може містити проєкції верхніх рівнів, структурні деталі, елементи проводки та елементи ландшафту. У більшості випадків план поверху представляє планування кімнат, включаючи функції та розміри, коридори, отвори, елементи інтер'єру та меблі, такі як каміни та раковини. Чим більше план поверху включає в себе деталей, тим більше він містить зіставних архітектурних елементів і стає складнішим з точки зору розкриття відповідних особливостей.

Вибір типу будівлі та деталей її поверхового плану набувають значення для підготовки набору даних. Коли масштаб будівлі збільшується, кількість архітектурних елементів і, отже, деталей її поверхового плану також збільшується, незалежно від того, чи це концептуальне чи фактичне креслення. Набір даних плану поверху створюється з растрових зображень плану поверху, що представляють стіни, анотації, вікна, двері та меблі

Хоча вся графічна інформація, наведена вище, важлива для представлення будинку в 2D-середовищі, усі виступи, меблі та сходи не враховуються, щоб зберегти те, що важливо для створення 3D-моделі: структурні елементи та отвори.

Цілісність структурних елементів можна оцінити лише на планах поверхів, і це єдине джерело, яке можна використовувати для створення структурних елементів у 3D-середовищі. Індикація цих елементів може бути досягнута лише шляхом їх геометричного позначення. Внутрішні та зовнішні отвори також можна витягти з плану поверху, однак вони потребують людського втручання для 3D-трансформації, оскільки точне розташування та розміри неможливо розгадати. Враховуючи цю ситуацію, необхідно розробити систему маркування для визначення основних архітектурних елементів на плані поверху.

Ця система оцінювання має базуватись на представленні очікуваного результату процесу навчання. Тому його слід підготовлювати шляхом введення нових зображень, які містять і відображають точну збережену геометричну інформацію креслень. Такі зображення називають набором міток, вони використовуються для сегментування стін і отворів у растровому зображенні плані поверху.

*Набір міток плану поверху* включає 2D зображення, на яких стіни позначено суцільними жирними лініями, що проходять через зовнішні отвори, а внутрішні отвори рознесені, оскільки внутрішні отвори можна отримати лише з планів поверхів, а специфікації зовнішніх отворів можна легко отримати з креслень висоти з посиланням до планів поверхів.

*Набір даних висоти.* Висота — це плоске зображення виду, який видно з одного боку будівлі. Оскільки будівля зазвичай не має простого прямокутного плану, креслення рельєфу представляє собою проекцію всіх частин будівлі, які видно з певного напрямку зі сплющеною

перспективою. Основна характеристика фасаду полягає в тому, що він не включає жодної глибини, однак використання планів поверхів і фасадів одночасно може подолати цю проблему. Загалом висоти генеруються для чотирьох напрямків, таких як схід, захід, північ і південь, щоб надати вичерпну інформацію про зовнішній вигляд будівлі. Фасади можуть містити різний рівень деталей залежно від зони використання, як на планах поверхів, але прості креслення фасадів показують анотації, такі як опорні рівні та розміри, зовнішні стіни, отвори, дахи та виступи.

Є дві основні проблеми у «висотах»:

- Зміст креслення рельєфу може збільшуватися зі збільшенням масштабу будівлі. Прості креслення одноповерхових і двоповерхових житлових будинків складаються як набір даних висоти, щоб бути узгодженим із набором даних плану поверху та мінімізувати цю проблему.

- Фасади можуть складатися з різних стилів представлення на відміну від планів.

Використання відтінків, тіней, ландшафтних елементів, перил, матеріалів, різних типів вікон і дверей може відрізнитися від дизайнера до дизайнера. У результаті стає важко визначити закономірність. Ця ситуація вимагає мати більший набір даних про висоту, ніж план поверху, щоб належним чином вивчити об'єкти. Набір даних висот створюється з растрових простих креслень висот зі стінами, отворами, анотаціями, дахами та виступами. Уже згадувалося, що стіни та отвори є головним поняттям для реконструкції 3D-моделі, а найточніші зовнішні отвори доступні лише в наборі даних висоти. Подібно до позначення набору даних плану поверху, висоти позначаються відповідно, щоб створити відповідний набір міток. Знову ж таки, основна ідея набору міток висоти полягає в представленні очікуваного результату процесу навчання. Набір міток висоти готується шляхом

позначення важливої геометричної інформації на кресленнях. Для набору позначено лише стіни та отвори, оскільки висоти є додатковим джерелом даних для набору даних плану поверху. Отже, стіни позначаються суцільно чорним, а отвори – суцільно білим для позначення набору даних висоти.

*Покращення даних.* Не завжди можливо створити набір даних із тисячами зображень, як у наборах даних ImageNet або MSCOCO. Це потребує багато обчислень, часу та праці. Якщо немає можливості зібрати більше екземплярів, дані можна покращити, щоб розширити доступний набір даних. Покращення даних збільшує вибірки даних зображення за допомогою основних перетворень, таких як обертання та масштабування. Ідея покращення даних, очевидно, полягає в тому, щоб мати збільшений набір даних, щоб мати можливість вивчити більш узагальнений шаблон на зображеннях. Якщо немає відповідного набору даних з точки зору розміру та однорідності, процес вивчення шаблонів і особливостей може виявитися невдалим, що призведе до *overfitting* або *underfitting*. У той час як *overfitting* означає, що набір даних вивчається повністю, *underfitting* є результатом неадекватного вивчення набору даних. У випадках *overfitting* навчання фактично не відбувається, а відбувається запам'ятовування. Таким чином, система погано працює на будь-якому новому зразку. Подібним чином новий (не визначений) зразок не розпізнається як *underfitting*, оскільки система не може дізнатися навіть сам набір даних. Потенціал покращення даних здебільшого запобігає цим проблемам шляхом збільшення вибірки даних і забезпечує більш надійний процес розпізнавання. Збільшуючи набір даних, важливо враховувати побічні ефекти. Наприклад, якщо кожне зображення в наборі даних буде центровано у кадрі, це стане відмінною рисою набору даних.



Таким чином, об'єкт на зображенні розпізнається відповідно до розташування, а не кутів і країв. Ось чому слід застосовувати перетворення, вирівняні за кутами, щоб урізноманітнити набір даних.

Як інший приклад, горизонтальне або вертикальне дзеркальне відображення слід передбачити заздалегідь, оскільки це може бути нерелевантним для зображення, яке буде перевернуто.

2) *Information extraction* – методика навчання та тестування нейронної мережі.

Серед методів виділення ознак застосування семантичної сегментації є більш зручним з огляду на цілі цієї дисертації. Архітектурні креслення містять надмірну кількість інформації. Важко слідувати шаблону та побудувати узагальнений підхід на основі правил для спеціального набору даних, як у підходах на основі форми та на основі зв'язків.

На відміну від цих алгоритмів, семантична сегментація не потребує попередньо визначеного набору правил. Натомість вона (семантична сегментація) сприймає зображення на рівні пікселів і призначає кожен піксель класу об'єктів відповідно до шаблону, створеного шляхом ітерації набору даних.

Якщо розглядати на прикладі будівельної галузі, то відповідно крім того що об'єкти зображення мають назву «стіна» та «вікно», ці об'єкти також геометрично представлені в алгоритмах семантичної сегментації. Ці представлення можуть належати до двох категорій («стіна» чи «не стіна») або до кількох класів («стіна», «вікно», «двері»), що має бути визначено заздалегідь.

У цьому дослідженні використовуються дві категорії, оскільки відтворення 3D-моделі базується на структурних елементах і отворах.

Оскільки використання лише категоріальної інформації про архітектурні елементи недостатньо для створення 3D-моделі, повинні бути геометрії (наприклад, прямокутні, круглі, трикутні,

неортогональні форми), які потрібно трансформувати в 3D-середовище.

Таким чином, семантично виділені архітектурні елементи стануть наріжним каменем для створення 3D-моделі. Семантична сегментація завжди була однією з важливих концепцій реконструкції 3D-моделі. Багато програм потребують точних, ефективних і надійних методів сегментації для оцінки візуальних елементів і прийняття рішень у реальному часі.

Ці методи можна класифікувати як традиційні підходи машинного навчання та застосування нейронної мережі.

Традиційні методи вимагають етапу попередньої обробки для збору функцій для прогнозування.

З іншого боку, штучні нейронні мережі (ANN) приймають необроблені вхідні дані та обробляють їх у ієрархічній біологічній структурі нейронів, щоб дізнатися параметри заданих вхідних даних.

Згідно зі спостереженнями, CNN є найбільш придатними та узгодженими моделями завдяки прямій сумісності з сегментацією семантичного зображення.

Розподіл ваги, можливість вилучення як низькорівневих, так і високорівневих особливостей у зображенні, використання меншої кількості параметрів, можливість налаштування для різних наборів даних, можливість трансформації в інші моделі, можливість використання попередньо навчених вагових коефіцієнтів моделі є суттєвими причинами для застосування CNN. CNN — це, по суті, нейронна мережа, яка має вагомість і упередження, які можна вивчити. Кожен нейрон у системі отримує певний вхідний об'єм, кодує його характеристики та перетворюється на інший нейрон за допомогою диференційованої функції.

На відміну від звичайних нейронних мереж, які приймають одновимірний вектор і використовують повністю пов'язані шари, CNN працюють з 3D-векторами розміром  $h, x, w, d$ :

де

$h$  — висота зображення,

$x$  — глибина зображення,

$w$  — ширина зображення,

$d$  — колірний канал.

CNN можуть служити як для завдань класифікації, так і для семантичної сегментації. Якщо завдання полягає в класифікації, реалізується кодування з повним зв'язним шаром, який в основному є звичайною нейронною мережею. Якщо потрібна семантична сегментація, застосовується декодування після кодування. CNN для семантичної сегментації складається з серії шарів, які можна назвати «Рівень згортки», «Активаційний рівень», «Рівень зменшення дискретизації» та «Рівень підвищення дискретизації». Ці рівні можна змінювати в залежності від завдання та набору даних. Рівень згортки містить фільтр/ядро для аналізу вхідного зображення, доки не будуть охоплені всі значення пікселів. Значення пікселів фільтра та зображення множаться під час згортки. Після цієї лінійної операції вводиться рівень активації, який зазвичай містить різні функції, такі як випрямлені лінійні одиниці (Rectified Linear Units) та сигмоїдна функція, щоб додати системі нелінійність. Рівень зменшення дискретизації головним чином допомагає усунути витрати на обчислення шляхом зменшення розмірності карти активації. Рівень підвищення дискретизації покращує результат попереднього рівня активації з низької роздільної здатності на більш високу. Таким чином, розмір передбаченого зображення буде таким самим, як розмір вхідного.

Усі ці кроки налаштовані відповідним чином для досягнення добре виконаної семантичної сегментації, продуктивність якої оптимізовано функцією вартості. Основна мета полягає в тому, щоб максимально мінімізувати функцію витрат шляхом оптимізації набору даних і структури CNN.

Функція вартості оцінює прогнози класу міток для кожного пікселя на зображенні та усереднює для всіх пікселів.

Двійкова перехресна ентропія використовується як функція вартості цього дослідження, оскільки завдання полягає в бінарній семантичній сегментації.

### 3) *3D model construction* – концепція побудови 3D моделі.

Після семантичної сегментації архітектурних креслень за допомогою архітектур CNN у двовимірному піксельному середовищі стає можливим 3D-перетворення для отримання 3D-реконструйованих моделей. Перед векторизацією архітектурних компонентів у семантично сегментованому зображенні виконується етап постобробки для видалення шумів із зображень.

Існують різні підходи до усунення явних помилок знайденої сегментації.

Векторизація прогнозованого зображення в першу чергу спирається на морфологічні перетворення для покращення зображення. Пізніше покращені зображення контуруються. Цей підхід в основному використовується завдяки простим математичним перетворенням і легкості застосування.

Морфологічні перетворення зазвичай використовуються для виявлення країв, видалення шуму, покращення зображення та сегментації зображення в програмах обробки зображень. Перетворення, яке застосовується до 2D зображень або зображень у градаціях сірого, включає морфологічний оператор на основі структуруючого елемента.

Структуруючий елемент — це в основному ядро, яке містить шаблон, заданий заздалегідь відносно початкової точки, і ковзає навколо зображення. Якщо структуруючий елемент і базове зображення відповідають вимогам, визначеним оператором, піксель зображення під джерелом структуруючого елемента встановлюється на попередньо визначене значення. Морфологічні оператори дозволяють пікселям перетинатися, об'єднуватися та/або заповнюватися.

Ерозія та розширення є основними операторами, але існують варіанти, такі як відкриття, закриття та розрідження. Оператор ерозії звужує межі пікселів переднього плану, тому області переднього плану стають меншими, а отвори в цих областях стають більшими. Навпаки, оператор розширення розширює пікселі переднього плану, і вони стають більшими, а отвори стають меншими.

Оператор закриття, який є оператором розширенням з наступною ерозією, використовується як для плану поверху, так і для передбачення висоти.

Для оптимізації ресурсів, що використовуються одночасно, було взято структуруючий елемент прямокутної форми розміром 3x3 для прогнозування плану поверху.

Будь-який інший спеціальний структуруючий елемент для отворів не використовується, оскільки в сегментах плану поверху є лише стіни. Для сегментації висоти структуруючий елемент складається з прямокутної форми на основі сітки розмірів висоти стіни та ширини для позначення загальної стіни.

Для екстракції достатньо структурного елемента прямокутної форми розміром 3x3.

*Контуриг.* Будь-яке чорно-біле зображення можна обчислити за допомогою «0» і «1», що представляють інтенсивність кольору. Ці 0 і 1 відповідають один одному, щоб можна було побудувати зв'язок.

По-перше, розмір зображення в пікселях приймається як початкова точка у верхньому лівому куті. Алгоритм відпрацьовує крок за кроком спочатку зверху вниз, потім зліва направо. Під час обробки кожного пікселя, якщо він стикається з «1», він зараховує цей «1» як дочірню точку межі та дає нове значення. Процес продовжується з тими самими кроками, і нарешті точки з однаковими значеннями розпізнаються та з'єднуються для створення контуру. Контуровання морфологічно трансформованого плану поверху та передбачення висоти є простим завданням, оскільки вони є 2д зображеннями.

Відтворення 3D. Контури містять інформацію про 2D точку на зображенні. Таким чином, ці точки забезпечують середовище перетворення 2D у 3D.

Процес перетворення складається з шести кроків:

- По-перше, кожен із контурів висоти вирівнюється відповідно до плану поверху на північ, схід, південь і захід. На даний момент немає рішення для прогнозування напрямків виходячи з архітектурного креслення, тому користувач повинен заздалегідь вказати висоти.
- Після вирівнювання висоти слід масштабувати відповідно до довжини країв плану поверху, оскільки висоти можуть не бути такими, що відповідають розміру зображення. Вирівнювання також надає інформацію про висоту. Це важливий крок для автоматизованого потоку, оскільки в більшості попередніх досліджень висота в основному визначається як значення за замовчуванням.
- Потім інформація про висоту використовується для екструдовання контурів плану поверху через вісь z.
- Нарешті, вирівняні висоти та екструдована маса стіни об'єднуються.
- Контури отворів на підвищеннях видавлюються і перетинаються з масою.
- Таким чином, області, що перетинаються, вирізаються з 3D об'єкту. Кінцевий результат моделі включає лише підлогу, зовнішні стіни з

отворами у вигляді порожнечі, внутрішні стіни та колони, якщо вони є.

*Розглянемо методи інформаційної технології розпізнавання графічної інформації на основі нейронної мережі:*

Принцип роботи системи полягає в обробці наборів даних, які включають два основні типи файлів креслень. Перший тип включає креслення поверхні, що являє собою геометричні форми, розташовані на плоскій поверхні. Включно з інформацією про розташування об'єктів, їх форму та розмір. Другий тип містить креслення висот, що представляє простір у трьох вимірах, ілюструючи висоту або глибину об'єктів відносно певної початкової точки.

Першим ключовим етапом обробки введених наборів даних є процес анотації або «лейблінгу», який полягає в мануальному введенні інформації, необхідної для тренування нейронних мереж. Анотація даних є важливою складовою процесу машинного навчання, оскільки вона дає моделям навчання контекст, необхідний для інтерпретації вхідних даних.

Після завершення процесу анотації вхідні дані надсилаються на обробку до нейронних мереж. Для цієї задачі використовуються різні типи нейронних мереж, включаючи глибокі згорткові нейронні мережі (DCNN) та багат шарові перцептрони (MLP). Ці нейронні мережі, використовуючи великі обсяги вхідних даних, спроможні виявляти та вивчати складні шаблони і залежності.

Сучасні технології баз даних, такі як SQL (в конкретному випадку PostgreSQL) використовуються для зберігання стану нейронних мереж, що дозволяє забезпечити стійкість моделі та її ітеративне вдосконалення. Зберігаючи навчальний стан кожної нейронної мережі, система має можливість проводити послідовні ітерації навчання на основі попередніх результатів.

Відповідно, процес ітеративного вдосконалення машинного навчання має рішуче значення в галузі комп'ютерного зору та моделювання, оскільки

він дозволяє забезпечити поступове поліпшення якості вивчення моделей. Після кожної ітерації, нейронні мережі використовують усі доступні дані, щоб покращити точність вивчення моделі.

Результатом цього вивчення є генерація векторизованих представлень поверхні та висотної діаграми. Ці представлення, які складаються з числових векторів, що описують відносне розташування та характеристики об'єктів на поверхні та висотної діаграми, стають вихідними даними для наступних етапів обробки.

З цих векторизованих даних, система створює агрегований набір даних, який описує стіни та висоти. Цей набір даних перетворюється у фінальний векторизований набір даних, що включає детальний опис стін та отворів.

В кінцевому підсумку, використовуючи високоефективні алгоритми комп'ютерного зору та машинного навчання, векторизовані дані конвертуються у 3D-модель.

На основі концептуальної моделі, методик і методів визначимо *генезу інформаційної технології розпізнавання графічної інформації на основі нейронної мережі* (рис. 3.2):

У рамках дисертаційного дослідження під *генезою інформаційної технології розпізнавання графічної інформації* будемо розуміти процес утворення або перетворення / трансформації  $2D \rightarrow 3D$  на основі комплексу методик та концепції, які застосовуються для розпізнавання графічної інформації та перетворення її із двовекторного простору у тривекторний.

Отже, враховуючи авторське визначення генези інформаційної технології розпізнавання графічної інформації на основі нейронної мережі та дані рисунка 3.2. опишемо генезу інформаційної технології як скупність етапів процесу розпізнавання графічної інформації та її трансформації у тривимірний простір із застосуванням визначених методик і методів на кожному етапі.



<b>ГЕНЕЗА ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ РОЗПІЗНАВАННЯ ГРАФІЧНОЇ ІНФОРМАЦІЇ</b>			
<b>ЕТАПИ</b>	<b>Data preparation</b>	<b>Information extraction</b>	<b>3D model construction</b>
<b>МЕТОДИКИ</b>	<b>підготовки даних для навчання нейронної мережі</b>	<b>навчання та тестування нейронної мережі</b>	<b>концепція побудови 3D моделі</b>
<b>МЕТОДИ</b>	<ul style="list-style-type: none"> <li>✓ машинного навчання,</li> <li>✓ нейронних мереж, включаючи глибокі згорткові нейронні мережі (DCNN) та багат шарові перцептрони (MLP),</li> <li>✓ послідовних ітерацій навчання на основі попередніх результатів.</li> </ul>	<ul style="list-style-type: none"> <li>✓ методи комп'ютерного зору та моделювання,</li> <li>✓ методи генерації векторизованих представлень</li> </ul>	<ul style="list-style-type: none"> <li>✓ методи алгоритмів комп'ютерного зору</li> <li>✓ машинного навчання.</li> </ul>

**Рисунок 3.2. Генеза інформаційної технології розпізнавання графічної інформації на основі нейронної мережі**

*Джерело: побудовано автором*

На першому етапі (Data preparation) виконується методика підготовки даних для навчання нейронної мережі із використанням методів машинного навчання, нейронних мереж, включаючи глибокі згорткові нейронні мережі (DCNN) та багат шарові перцептрони (MLP) та послідовних ітерацій навчання на основі попередніх результатів.

На другому етапі (Information extraction), за допомогою методики навчання та тестування нейронної мережі на основі методів комп'ютерного зору та моделювання і методів генерації векторизованих представлень,

відбувається процес вилучення інформації з графічних даних, які отримані на попередньому етапі.

На третьому етапі (3D model construction) здійснюється побудова 3D моделі, за допомогою методів комп'ютерного зору та моделювання, а також методів машинного навчання.

В результаті вище описаних дій відбувається процес розпізнавання та перетворення графічних примітивів із двовимірного у тривимірний простір.

### **3.2. Контекстна модель інформаційної технології розпізнавання графічної інформації на основі нейронної мережі**

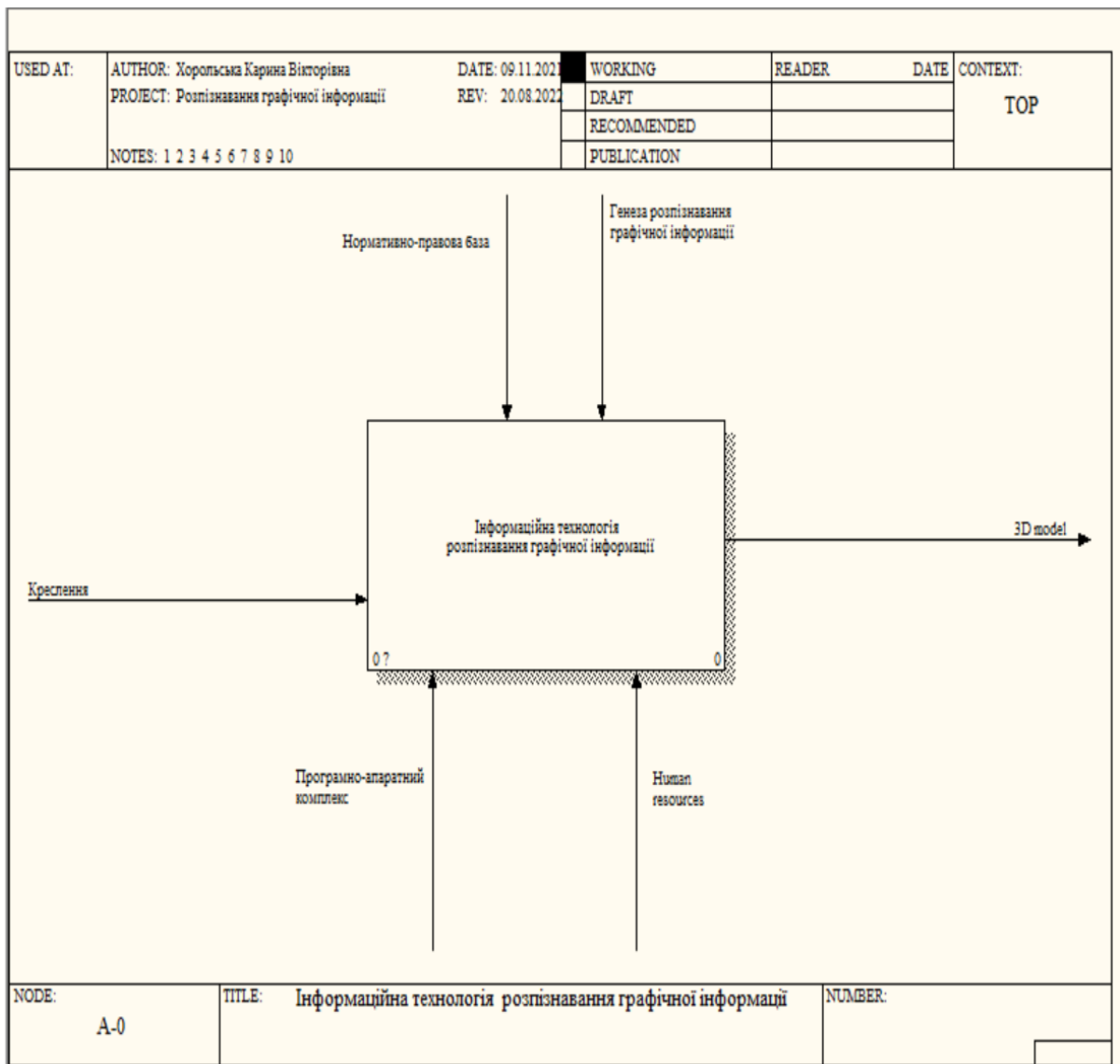
Одним з представлень інформаційної технології розпізнавання графічної інформації на основі нейронної мережі є розробка контекстної моделі, що являє собою декомпозицію декількох рівнів. На рисунку 3.3 показано вхідні компоненти та ключові елементи наповнення роботи інформаційної технології, яка розпізнає зображення 2D формату, при вході отримується креслення і в процесі обробки використання моделі видає 3D модель, діяльність технології обов'язково враховує також інші вхідні дані.

Якщо розглядати дану нотацію з точки зору вхідної (Input, Control, Mechanism) та вихідної (Output) інформації, то дану модель можна розглянути у двох варіантах: 1 етап - Навчання системи; 2 етап – Використання системи.

Перший та другий етапи різняться надавачем первинної інформації для входу в систему та типом первинної інформації.

1) Input: первинні дані надаються системі у вигляді креслення, що представляє собою цифровий формат зображень зовнішніх частин об'єкту, що потребує розпізнавання та проектування у 3D форматі.

На першому етапі під час навчання системи дані до системи надаються розробником системи (Database manager), що є частиною людського ресурсу (Human resources).



**Рисунок 3.3. Контекстна модель «Інформаційна технологія розпізнавання графічної інформації»**

*Джерело: побудовано автором*

На другому етапі, тобто під час використання системи дані до системи надає користувач, що є частиною людського ресурсу (Human resources).

## 2) Вхідна інформація:

1. Нормативно-правова база - закони України, які визначають загальні засади оброблення інформації як інформаційного продукту та інформаційного ресурсу за допомогою інформаційно-інтелектуальних технологій.

2. Генеза розпізнання графічної – комплекс методик та концепції, які застосовуються для розпізнавання графічної інформації та перетворення її із двовекторного простору у тривекторний.

3. Програмно-апаратний комплекс – набір технічних та програмних засобів, що працюють спільно для виконання завдань пов'язаних із функціонуванням нейромережі.

4. Human resources:

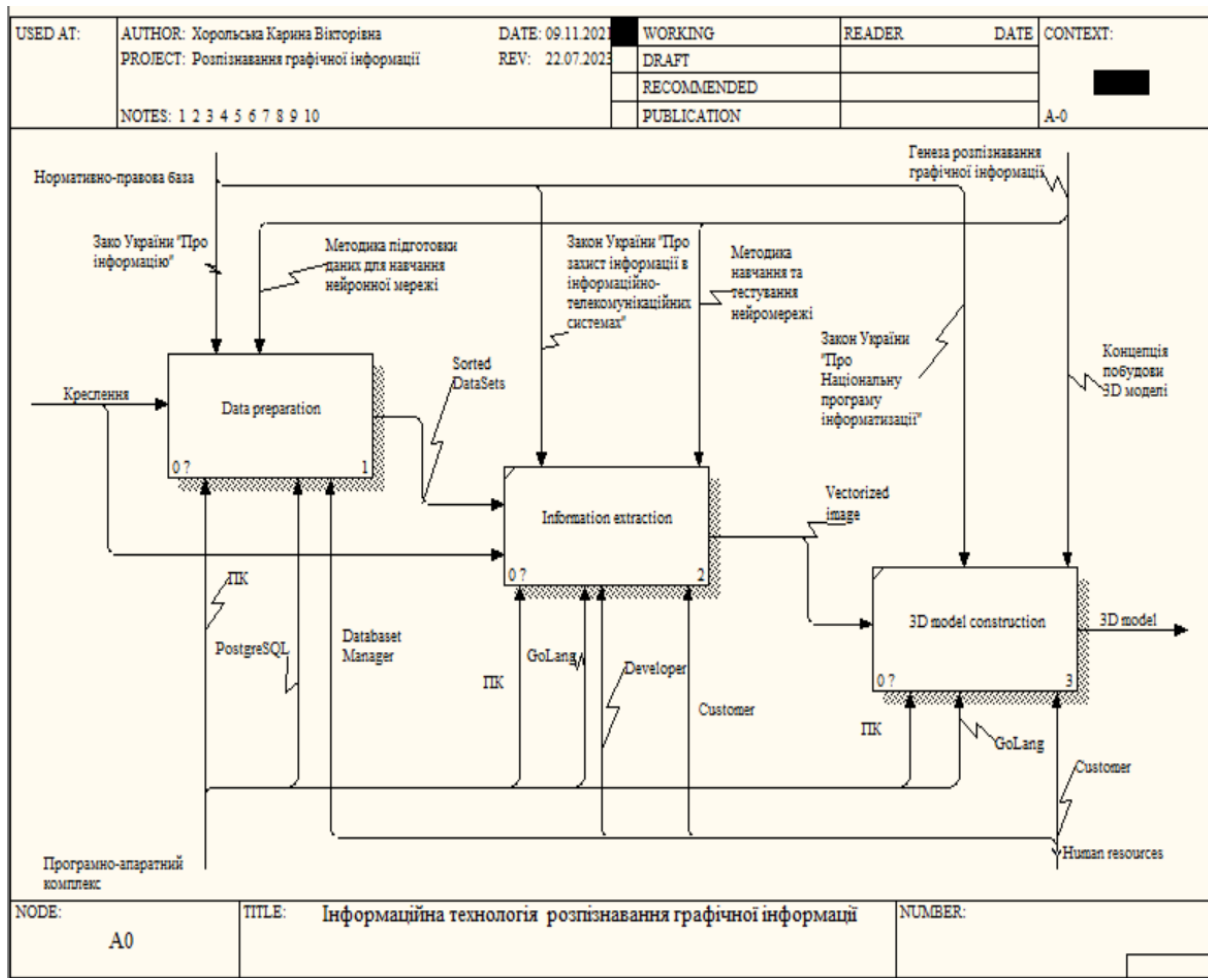
а. Висококваліфіковані представники ІТ професії, які безпосередньо займаються навчанням нейромережі коректного розпізнавання графічної інформації, що буде використана для моделювання 3D образів (на першому етапі – Навчання системи);

б. Користувачі (споживачі), які використовуватимуть дану інформаційну технологію у своїй професійній діяльності (на другому етапі – Використання системи).

3) Output: 3D model (3D модель) - результат оброблення та перетворення інформації із двовекторного простору у тривекторний нейроною мережею.

Після врахування і обробки кожного вхідного елемента технологія генерує та обробляє отримані дані та на виході видає готову 3D модель отриманого на вході креслення.

Модель інформаційної технології в діаграмі першої декомпозиції інформаційної технології розпізнавання графічної інформації має три ключові блоки (рис. 3.4): Data preparation, Information extraction, 3D model construction.



**Рисунок 3.4. Діаграма першої декомпозиції «Інформаційна технологія розпізнавання графічної інформації»**

*Джерело: побудовано автором*

Повну деталізацію діаграми першої декомпозиції наведено в таблицях Додатку А.

Дана модель складається із трьох складових:

- Підготовка даних (Data preparation) - розподілення вхідної інформації як первинний етап в навчанні нейронної мережі: перетворення технічного креслення у зображення двовимірного простору. Процес навчання (з подальшим тестуванням) нейронної мережі будується на основі сприйняття зображення на рівні пікселя і призначення кожному пікселю класу об'єктів відповідно до шаблону, який нейронна мережа створює шляхом ітерації набору даних. Для

виділення архітектурних елементів використовується CNN (Convolutional Neural Networks).

- Оброблення інформації (Information extraction) – розпізнавання на основі алгоритмів CNN введеного користувачем технічного креслення (графічної інформації) та перетворення його з двовекторного формату у тривекторний для подальшої побудови 3D моделі.
- Побудова 3D моделі (3D model construction) - моделювання 3D образу розпізнаної графічної інформації на основі отриманих даних від нейромережі.

Ці компоненти пов'язані між собою вхідними та вихідними даними.

Охарактеризуємо компоненти даної нотації:

#### 1) *Data preparation:*

##### *Input:*

Креслення - фронтальний вигляд об'єкта, бокове зображення, зображення ззаду, план зверху.

##### *Control:*

- Закон України «Про інформацію» - Закон, що регулює відносини щодо створення, збирання, одержання, зберігання, використання, поширення, охорони, захисту інформації [92].
- Методика підготовки даних для навчання нейронної мережі - комплекс правил та рекомендацій щодо підготовки первинних даних для навчання нейронної мережі.

##### *Mechanism:*

- ПК – мінімальні технічні вимоги до ПК:

CPU - Intel Core i7-12700K

GPU - GeForce RTX 3090 super

RAM - Corsair Vengeance LPX 32GB (2 X 16GB) 3600MHz DDR4 C18

Storage - Samsung 980 500GB NVMe Gen 3 SSD

## Motherboard - Gigabyte Z690 AORUS MASTER

- PostgreSQL – база даних.

Даний вибір обумовлений тим, що PostgreSQL є найліпшою базою даних з лінійки топ сегменту безкоштовних баз даних. Вигідною особливістю PostgreSQL є вдала, можливо найкраща, робота з бінарними даними. Вона чудово підходить для збереження зображень оперативного рівня. Отримання їх з бази значно швидше ніж читання з файлу, навіть за використання сучасних NVMe Gen 3 SSD.

- Database manager - технічний працівник, який маркерує первинні данні технічного креслення (графічної інформації), необхідний на етапі первинного навчання нейронної мережі.

### *Output:*

Sorted DataSets - відсортовані бази знань, які є первинними даними для навчання нейронної мережі.

## 2) *Information extraction:*

### *Input:*

- Креслення у вигляді цифрового формату зображень зовнішніх частин об'єкту, що потребує розпізнавання та проєктування у 3D форматі.
- Sorted DataSets - відсортовані бази знань, які є первинними даними для навчання нейронної мережі.

### *Control:*

- Закон України «Про електронні комунікації» - цей Закон регулює відносини у сфері захисту інформації в інформаційних, телекомунікаційних та інформаційно-телекомунікаційних системах (далі - система) [95].
- Методика навчання та тестування нейронної мережі - Комплекс програмних алгоритмів для формування бази даних та бази знань з метою розробки механізмів розпізнавання графічної інформації для переведення інформації у двовекторний простір. Тестування

нейронної мережі відбувається після попереднього процесу навчання з метою виявлення помилок та подальшим їх коригуванням.

*Mechanism:*

- Мінімальні технічні вимоги до ПК:
  - CPU - Intel Core i7-12700K
  - GPU - GeForce RTX 3090 super
  - RAM - Corsair Vengeance LPX 32GB (2 X 16GB) 3600MHz DDR4 C18
  - Storage - Samsung 980 500GB NVMe Gen 3 SSD
  - Motherboard - Gigabyte Z690 AORUS MASTER
- GoLang (Go) - мова програмування, яка забезпечує:
  - ✓ Відносно низький рівень входження в мову, що забезпечує максимальну «сумісність» з програмістами різного рівня кваліфікації
  - ✓ Має сучасний синтаксис та потужність низькорівневої мови програмування
  - ✓ Компілюється в машинний код, відповідно є максимально «нативною» та кросплатформенною за потреби.

IDE - GoLand

Вибір GoLand IDE обумовлений рядом фактів:

- Безкоштовна повна ліцензія для наукових цілей.
- Потужний інструментарій «з коробки» - встановив і використовуєш, система позбавляє користувача зайвих турбот.
- Великі можливості в базі.

Мінусом вибраної IDE є вибагливість до наявних ресурсів ПК.

- Developer – висококваліфікований розробник-програміст, який відповідає за реалізацію та програмний код нейронної мережі.
- Customer - кінцевий користувач технології розпізнавання графічної інформації.

*Output:*



Vectorized grid image - інтегроване 2D зображення з координатною трьохвимірною сіткою.

### 3) 3D model construction

#### *Input:*

Vectorized grid image - інтегроване 2D зображення з координатною трьохвимірною сіткою.

#### *Control:*

- Концепція побудови 3D моделі - система алгоритмів на основі яких нейронна мережа перетворює інформацію із двовекторного простору у тривекторний.

- Закон України «Про Національну програму інформатизації» - Національна програма інформатизації становить комплекс взаємопов'язаних окремих завдань (проектів) інформатизації, спрямованих на реалізацію державної політики та пріоритетних напрямів створення сучасної інформаційної інфраструктури України за рахунок концентрації та раціонального використання фінансових, матеріально-технічних та інших ресурсів, виробничого і науково-технічного потенціалу держави, а також координації діяльності державних органів, органів місцевого самоврядування, підприємств, установ, організацій усіх форм власності і громадян у сфері інформатизації [94].

#### *Mechanism:*

- Мінімальні технічні вимоги до ПК:
  - CPU - Intel Core i7-12700K
  - GPU - GeForce RTX 3090 super
  - RAM - Corsair Vengeance LPX 32GB (2 X 16GB) 3600MHz DDR4 C18
  - Storage - Samsung 980 500GB NVMe Gen 3 SSD
  - Motherboard - Gigabyte Z690 AORUS MASTER
- GoLang (Go) - мова програмування, яка забезпечує

- Відносно низький рівень входження в мову, що забезпечує максимальну «сумісність» з програмістами різного рівня кваліфікації
- Має сучасний синтаксис та потужність низькорівневої мови програмування
- Компілюється в машинний код, відповідно є максимально «нативною» та кросплатформенною за потреби.

IDE - GoLand

Вибір GoLand IDE обумовлений рядом фактів:

- Безкоштовна повна ліцензія для наукових цілей
- Потужний інструментарій «з коробки» - встановив і використовуєш, система позбавляє користувача зайвих турбот.
- Великі можливості в дебазі

Мінусом вибраної IDE є вибагливість до наявних ресурсів ПК.

- Customer - кінцевий користувач технології розпізнавання графічної інформації.

*Output:*

3D model - результат оброблення та перетворення інформації із двовекторного простору у тривекторний нейронною мережею.

Компоненти які представлені «Data preparation» пов'язані між собою вхідними та вихідними даними. Дана нотація складається із двох складових (рис. 3.5.):

- Surface blueprint - розподілені двовекторні креслення: фронтальний вигляд об'єкта, бокове зображення, зображення ззаду, план зверху.

- Side view - інтегровані 2D зображення: фронтальний вигляд об'єкта, бокове зображення, зображення ззаду.

Повну деталізацію діаграми другої декомпозиції наведено в таблицях Додатка А.

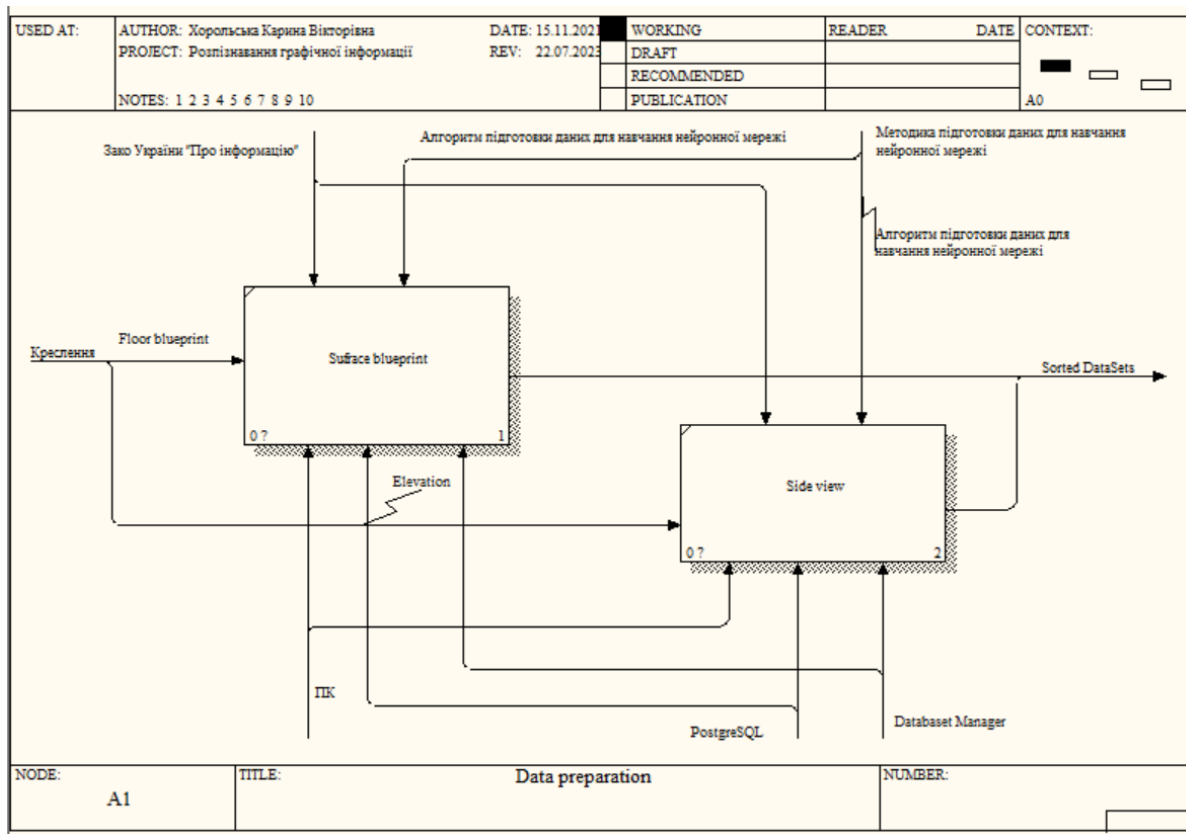


Рисунок 3.5. Деталізація Data preparation

Джерело: побудовано автором

### 1) Surface blueprint

#### Input:

Креслення (Floor blueprint) - фронтальний вигляд об'єкта, бокове зображення, зображення ззаду, план зверху.

#### Control:

- Закон України «Про інформацію» - цей Закон регулює відносини у сфері захисту інформації в інформаційних, телекомунікаційних та інформаційно-телекомунікаційних системах (далі - система) [92].

- Алгоритм підготовки даних для навчання нейронної мережі - комплекс програмних алгоритмів для формування бази даних та бази знань з метою розробки механізмів розпізнавання графічної інформації для переведення інформації у двовекторний простір. Тестування нейронної мережі відбувається після попереднього процесу навчання з метою виявлення похибок та подальшим їх коригуванням.

### *Mechanism:*

- Мінімальні технічні вимоги до ПК:

CPU - Intel Core i7-12700K

GPU - GeForce RTX 3090 super

RAM - Corsair Vengeance LPX 32GB (2 X 16GB) 3600MHz DDR4 C18

Storage - Samsung 980 500GB NVMe Gen 3 SSD

Motherboard - Gigabyte Z690 AORUS MASTER

- PostgreSQL – база даних. Найліпша продакшен базою даних з лінійки топ сегменту безкоштовних баз даних. Вигідною особливістю постгре є вдала, можливо найкраща, робота з бінарними даними. Вона чудово підходить для збереження зображень оперативного рівня. Отримання їх з бази значно швидше ніж читання з файлу, навіть за використання сучасних NVMe Gen 3 SSD.

- Database/DataSet Manager - технічний працівник, який маркерує первинні данні технічного креслення (графічної інформації), необхідний на етапі первинного навчання нейронної мережі.

### *Output:*

Educated DataSets - результат оброблення та перетворення інформації із двовекторного простору у тривекторний нейронною мережею: фронтальний вигляд об'єкта, бокове зображення, зображення ззаду, план зверху.

### *2) Side view:*

#### *Input:*

Креслення (Elevation) - 2D зображення.

#### *Control:*

- Закон України «Про інформацію» - цей Закон регулює відносини щодо створення, збирання, одержання, зберігання, використання, поширення, охорони, захисту інформації [92].

- Алгоритм підготовки даних для навчання нейронної мережі - комплекс програмних алгоритмів для формування бази даних та бази знань

з метою розробки механізмів розпізнавання графічної інформації для переведення інформації у двовекторний простір. Тестування нейронної мережі відбувається після попереднього процесу навчання з метою виявлення похибок та подальшим їх коригуванням.

*Mechanism:*

- Мінімальні технічні вимоги до ПК:

CPU - Intel Core i7-12700K

GPU - GeForce RTX 3090 super

RAM - Corsair Vengeance LPX 32GB (2 X 16GB) 3600MHz DDR4 C18

Storage - Samsung 980 500GB NVMe Gen 3 SSD

Motherboard - Gigabyte Z690 AORUS MASTER

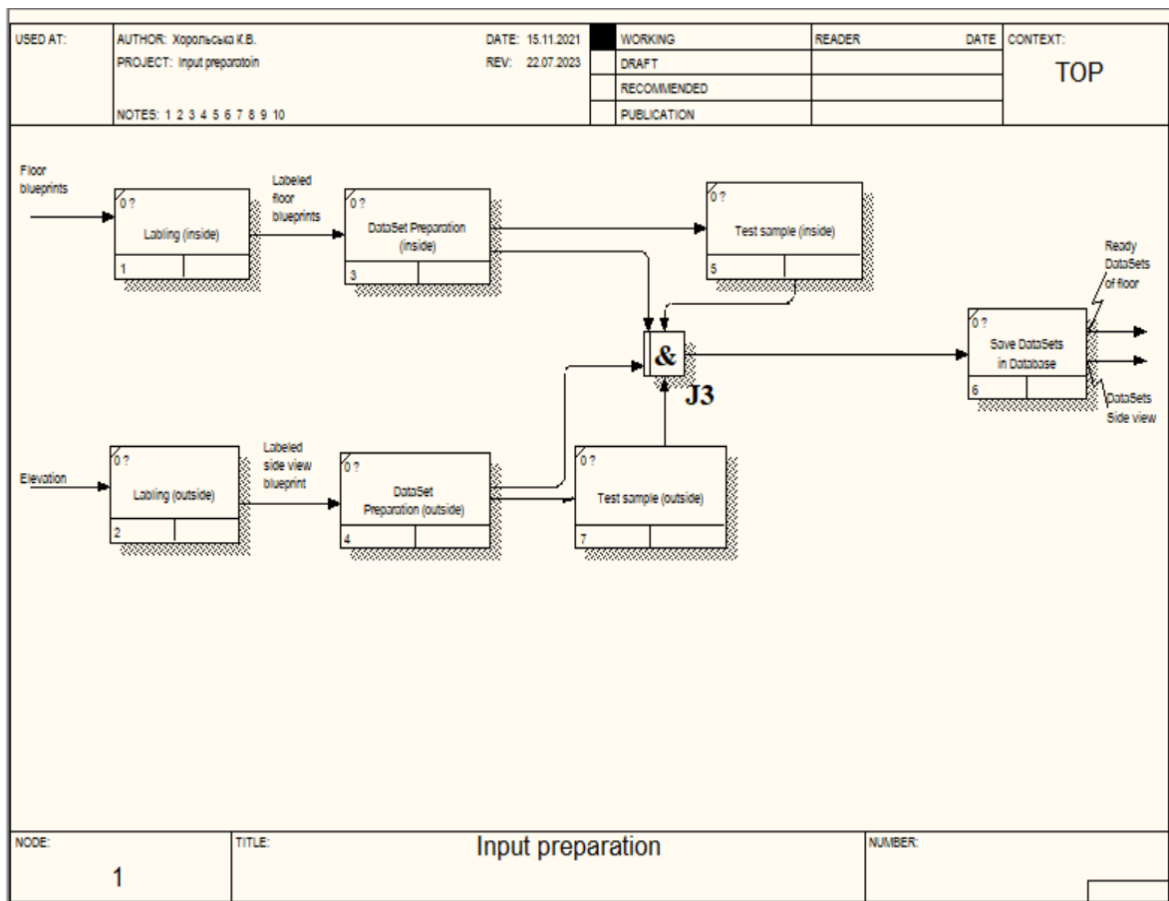
- PostgreSQL – база даних. Найліпша продакшен базою даних з лінійки топ сегменту безкоштовних баз даних. Вигідною особливістю постгре є вдала, можливо найкраща, робота з бінарними даними. Вона чудово підходить для збереження зображень оперативного рівня. Отримання їх з бази значно швидше ніж читання з файлу, навіть за використання сучасних NVMe Gen 3 SSD.

- Database/DataSet Manager - технічний працівник, який маркерує первинні данні технічного креслення (графічної інформації), необхідний на етапі первинного навчання нейронної мережі.

*Output:*

Educated DataSets - результат оброблення та перетворення 2D зображень із двовекторного простору у тривекторний нейронною мережею.

За допомогою контекстної діаграми IDEF3 графічно опишемо інформаційні потоки, взаємозв'язки між процесами обробки графічної інформації, що є частиною процесу підготовки вхідних даних для нейронної мережі (рис. 3.6)



**Рисунок 3.6. Деталізація процесу підготовки вхідних даних для нейронної мережі**

*Джерело: побудовано автором*

Модель IDEF3 декомпозиції процесу підготовки вхідних даних для нейронної мережі демонструє процес створення та підготовки наборів даних і тестових зразків для креслень поверхонь і фасадів, як для внутрішнього, так і зовнішнього маркування. Вона уточнює послідовність процесів та потік інформації, полегшуючи розуміння та аналіз процедури підготовки набору даних для зазначених компонентів.

Використовуючи можливості контекстної діаграми IDEF3 графічно опишемо інформаційні потоки, взаємозв'язки між процесами побудови 3D моделі (рис. 3.7)

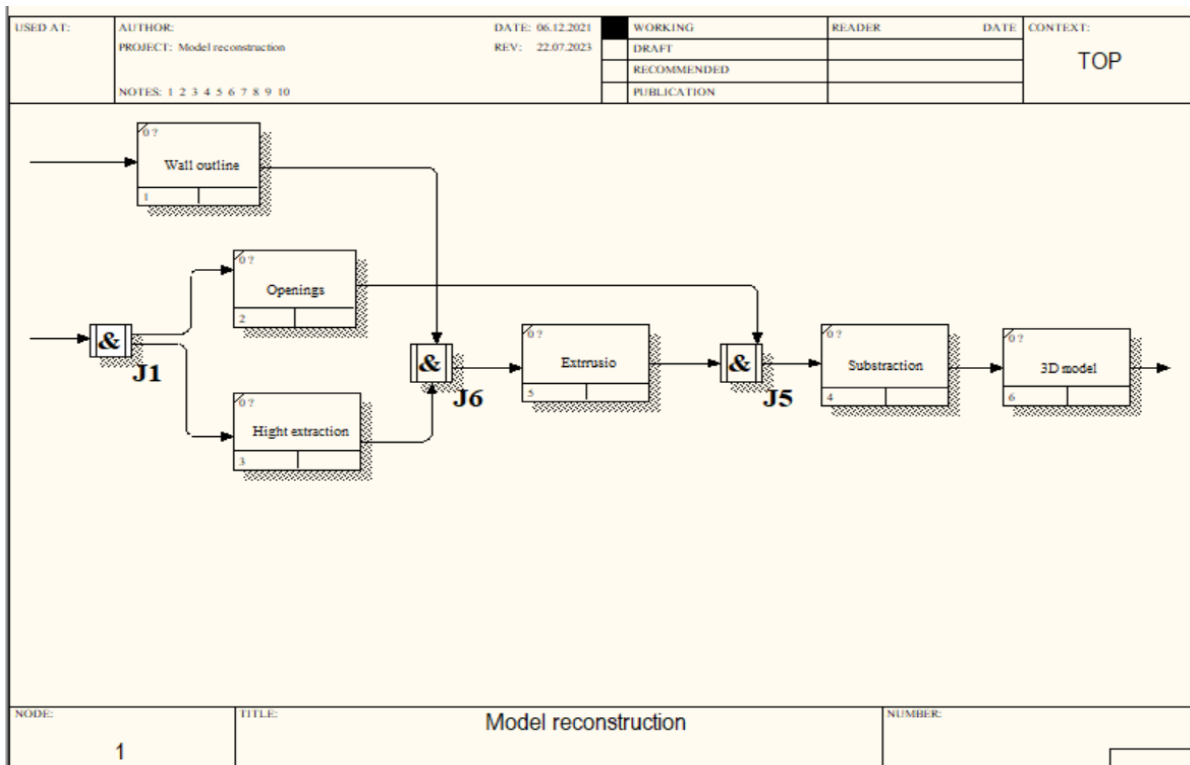


Рисунок 3.7. Деталізація процесу побудови 3D моделі

Джерело: побудовано автором

Представлена модель IDEF3 декомпозиції процесу підготовки вхідних даних для нейронної мережі ілюструє процес створення та підготовки наборів даних і тестових зразків для внутрішнього та зовнішнього маркування креслень поверхонь і фасадів. Модель включає кілька видів діяльності та інформаційних потоків (табл.3.1), як описано нижче:

Fan-out Junction і Fan-in Junction: ці з'єднувачі представляють потік інформації між різними діями. Fan-out Junction позначає початкову точку інформаційного потоку, тоді як Fan-in Junction вказує на кінцеву точку інформаційного потоку.

Surface blueprint: процес передбачає підготовку креслення поверхні.

Elevation: процес пов'язаний зі створенням проєктів висот.

Labling (inside): представляє процес маркування внутрішніх частин креслень площини об'єкта.

Labling (outside): представляє процес маркування зовнішніх частин креслень поверхонь та фасадів об'єкта.

DataSet preparation (inside): процес передбачає підготовку наборів даних для позначених внутрішніх частин.

DataSet preparation (outside): процес передбачає підготовку наборів даних для зовнішніх частин із мітками.

Test Sample (inside): процес створення тестових зразків для маркованих внутрішніх частин.

Test Sample (outside): процес створення тестових зразків для маркованих зовнішніх частин.

Precedence: анотації «Precedence» означають послідовний порядок, у якому мають відбуватися дії. Наприклад, «Surface blueprint» і «Elevation» повинні передувати «Labling (inside)» і «Labling (outside)», відповідно.

Asynchronous: анотація «Asynchronous» означає, що дії «DataSet preparation (inside)» не потрібно чекати завершення «Test Sample (inside)», перш ніж продовжити «Saving DataSets in DataStorage». Ці процеси можуть відбуватися незалежно.

Таблиця 3.1

Компоненти діаграми декомпозиції процесу підготовки вхідних даних для нейронної мережі

Назва компоненти	Характеристика	Назва компоненти
Fan-out Junction	інформаційних процесів	Fan-in Junction
Surface blueprint	Precedence	Labling (inside)
Elevation	Precedence	Labling (outside)
Labling (inside)	Precedence	DataSet preparation (inside)
Labling (outside)	Precedence	DataSet preparation (outside)
DataSet preparation (inside)	Precedence	Test Sample (inside)
DataSet preparation (outside)	Precedence	Test Sample (outside)
DataSet preparation (inside)	Asynchronous	Saving DataSets in DataStorage
DataSet preparation (outside)		
Test Sample (inside)		
Test Sample (outside)		

Джерело: побудовано автором



На рисунку 3.7 наведена IDEF3 модель, що демонструє процес створення тривимірної (3D) моделі об'єкта, включаючи різноманітні процеси: визначення контуру графічного об'єкта, створення отворів, вилучення елементів, видавлювання та віднімання частин для отримання остаточної 3D моделі. Дана модель також деталізує часові зв'язки між вказаними процесами (табл.3.2).

Таблиця 3.2

Компоненти діаграми декомпозиції процесу побудови 3D моделі

Назва компоненти	Характеристика інформаційних процесів	Назва компоненти
Fan-out Junction		Fan-in Junction
-	Precedence	Wall outline
-	Precedence	J1
J1	Asynchronous	Openings
		Hight extraction
Wall outline	Synchronous	Extrusion
Hight extraction		
Extrusion	Synchronous	Substraction
Openings		
Substraction	Precedence	3D model

*Джерело: побудовано автором*

Fan-out Junction та Fan-in Junction є вузлами, що забезпечують відображення потоку інформації між різними процесами. Fan-out Junction репрезентує точку початку потоку інформації, тоді як Fan-in Junction репрезентує точку кінця для потоку інформації.

Процес «Wall outline» визначає контур об'єкта.

«Openings» вказує на процес, що пов'язаний зі створенням отворів на площині об'єкта (на прикладі будівельної галузі: отвори у стіні, такі як вікна чи двері).

«High Extraction» представляє собою процес, який зумовлює вилучення специфічних елементів з контуру об'єкта, що може бути пов'язано зі створенням високорівневого представлення креслення.

«Extrusion» описує процес видавлювання або розширення певних частин контуру креслення з метою створення виступів або декоративних елементів.

«Substraction» представляє собою процес віднімання або видалення певних частин з контуру об'єкта з метою створення отворів або вирізів.

«3D model» є кінцевим результатом процесу, що представляє собою 3D-модель об'єкта, включаючи усі модифікації, що були здійснені під час дій «High Extraction», «Extrusion» та «Substraction».

Позначення «Precedence» та «Asynchronous/Synchronous» стосуються часових зв'язків між процесами. «Precedence» показує, що один процес має передувати іншому, тоді як «Asynchronous/Synchronous» вказує на можливість незалежної або синхронізованої реалізації процесів.

### **3.3. Архітектура програмного комплексу інформаційної технології розпізнавання графічної інформації на основі нейронної мережі**

Побудова архітектури програмного комплексу інформаційної технології розпізнавання графічної інформації на основі нейронної мережі є вирішальною для досягнення успіху у сфері комп'ютерного зору і обробки графічної інформації. Оптимальна архітектура не тільки забезпечує високу точність розпізнавання, але також гарантує ефективне використання обчислювальних ресурсів та відкриває можливість для широкого спектру застосувань цієї технології.

Спочатку опишемо Use-flow діаграму програмного додатку інформаційної технології розпізнавання графічної інформації на основі нейронної мережі, яка є важливим інструментом для розуміння, проектування та удосконалення досвіду користувача з додатком, що дає

можливість розробникові зосередитися на потребах користувача та забезпечити оптимальний та ефективний дизайн інтерфейсу додатку.

Діаграма відображена на рис. 3.8 відображає шлях користувача через різні екрани та функції додатку, допомагаючи розуміти його досвід та перспективу взаємодії з системою.

Розроблена Use-flow діаграма додатку складається з чотирьох основних блоків: Налаштування, Введення, Робота та Виведення.

*Налаштування:* У цьому блоці користувач налаштовує параметри конвертації. Це може включати вибір формату виводу, рівня деталізації 3D-моделі та інших параметрів, що впливають на процес конвертації.

*Введення:* На цьому етапі користувач вибирає файли для конвертації. Це можуть бути 2D-креслення, які відображають бічний та верхній види. Після вибору файлів система перевіряє їх на сумісність. Якщо файли не сумісні, система виводить повідомлення про помилку та повертає користувача до етапу введення файлів.

*Робота:* На цьому етапі система обробляє вхідні файли. Спочатку вона створює бінарні векторні зображення на основі 2D-креслень. Якщо виникає помилка, система виводить повідомлення про помилку та повертає користувача до етапу введення файлів. Після успішного створення векторних зображень система будує 3D-модель. Якщо виникає помилка під час побудови моделі, система також виводить повідомлення про помилку та повертає користувача до етапу введення файлів.

*Виведення:* На останньому етапі система зберігає результати до файлової системи. Якщо виникає помилка під час збереження, система виводить повідомлення про помилку та повертає користувача до етапу введення файлів.

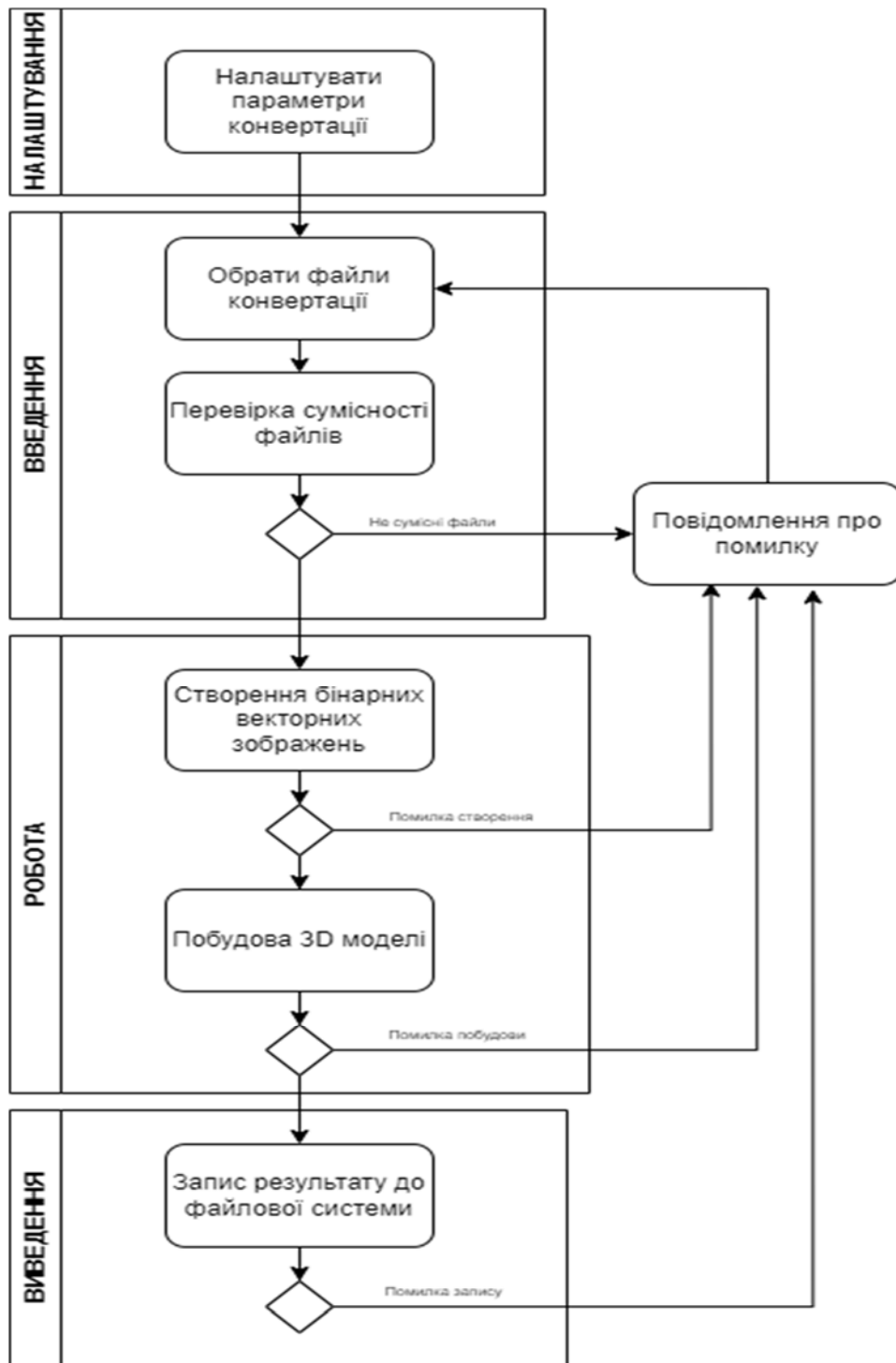


Рисунок 3.8. Use-flow діаграма додатку

Джерело: побудовано автором

Таким чином, ця діаграма відображає повний цикл роботи додатку, від налаштування параметрів і введення файлів до обробки даних і збереження результатів.

Окрім розуміння про повний цикл роботи додатку важливим є розуміння взаємодії між користувачами та системою. Use-case діаграма додатку інформаційної технології розпізнавання графічної інформації на основі нейронної мережі (рис. 3.9) додатку служить як основне відображення взаємодії між користувачами та системою, а також для аналізу, проектування та управління функціональністю додатку з орієнтацією на потреби користувача.

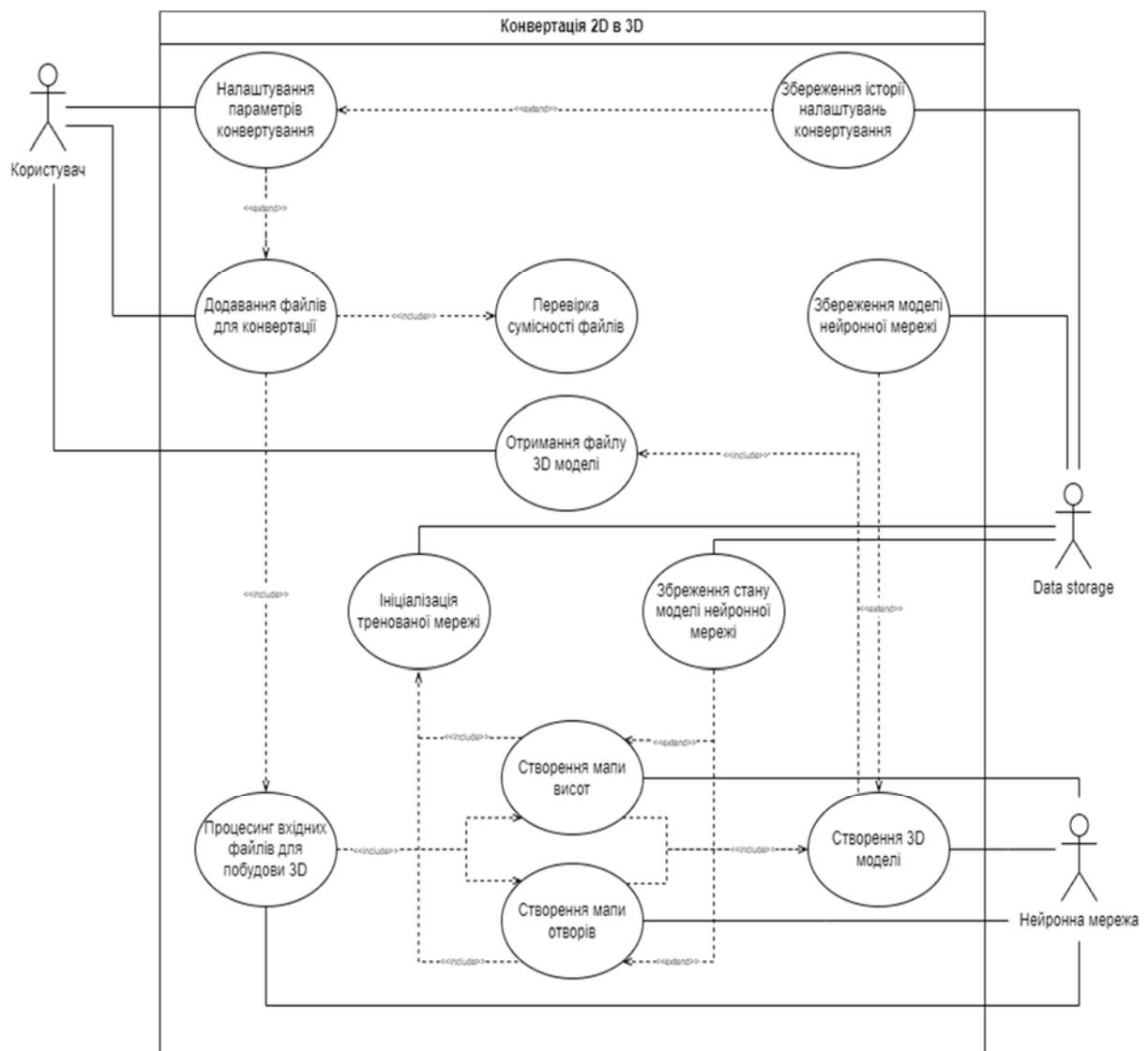


Рисунок 3.9. Use- case діаграма додатку

Джерело: побудовано автором

Опис use-case діаграми додатку інформаційної технології розпізнавання графічної інформації на основі нейронної мережі:

*Користувач* - основний актор, який взаємодіє з системою. Він може налаштовувати параметри конвертації, додавати файли для конвертації та отримувати файли 3D моделі.

*Налаштування параметрів конвертації* - це опціональний крок, який користувач може виконати перед додаванням файлів.

*Додавання файлів для конвертації* - на цьому етапі користувач додає файли, які потрібно конвертувати.

*Перевірка сумісності файлів* - система перевіряє, чи можуть бути конвертовані додані файли.

*Нейронна мережа* - це система, яка обробляє вхідні файли. Вона може проводити парсинг вхідних файлів для побудови 3D моделі, створювати мапи висот, створювати мапи отворів та створювати 3D моделі.

*Створення мап висот та мап отворів* - це процес, який виконує нейронна мережа після обробки вхідних файлів.

*Створення 3D моделі* - після створення мап висот та мап отворів, нейронна мережа створює 3D модель.

*Отримання файлу 3D моделі* - після створення 3D моделі, користувач може отримати файл 3D моделі.

Розробка логічної моделі системи у вигляді діаграми класів (рис. 3.10), яка відображає різні взаємозв'язки між окремими елементами, такими як об'єкти і підсистеми з атрибутами і операціями, а також описує їх внутрішню структуру і відношень [91].

Розроблена діаграма класів додатку інформаційної технології розпізнавання графічної інформації на основі нейронної мережі складається із 6 основних *MainWindowView*, *DevWindowView*, *MainWindowViewModel*, *DevWindowViewModel*, *MainWindowModel*, *DevWindowModel*.

Ця структура демонструє взаємодію користувача з системою та роль нейронної мережі у процесі обробки графічної інформації для створення 3D

моделей. Користувач має можливість налаштувати параметри конвертації та додавати файли для обробки, а система, в свою чергу, забезпечує розпізнавання та конвертацію графічної інформації на основі нейронної мережі.

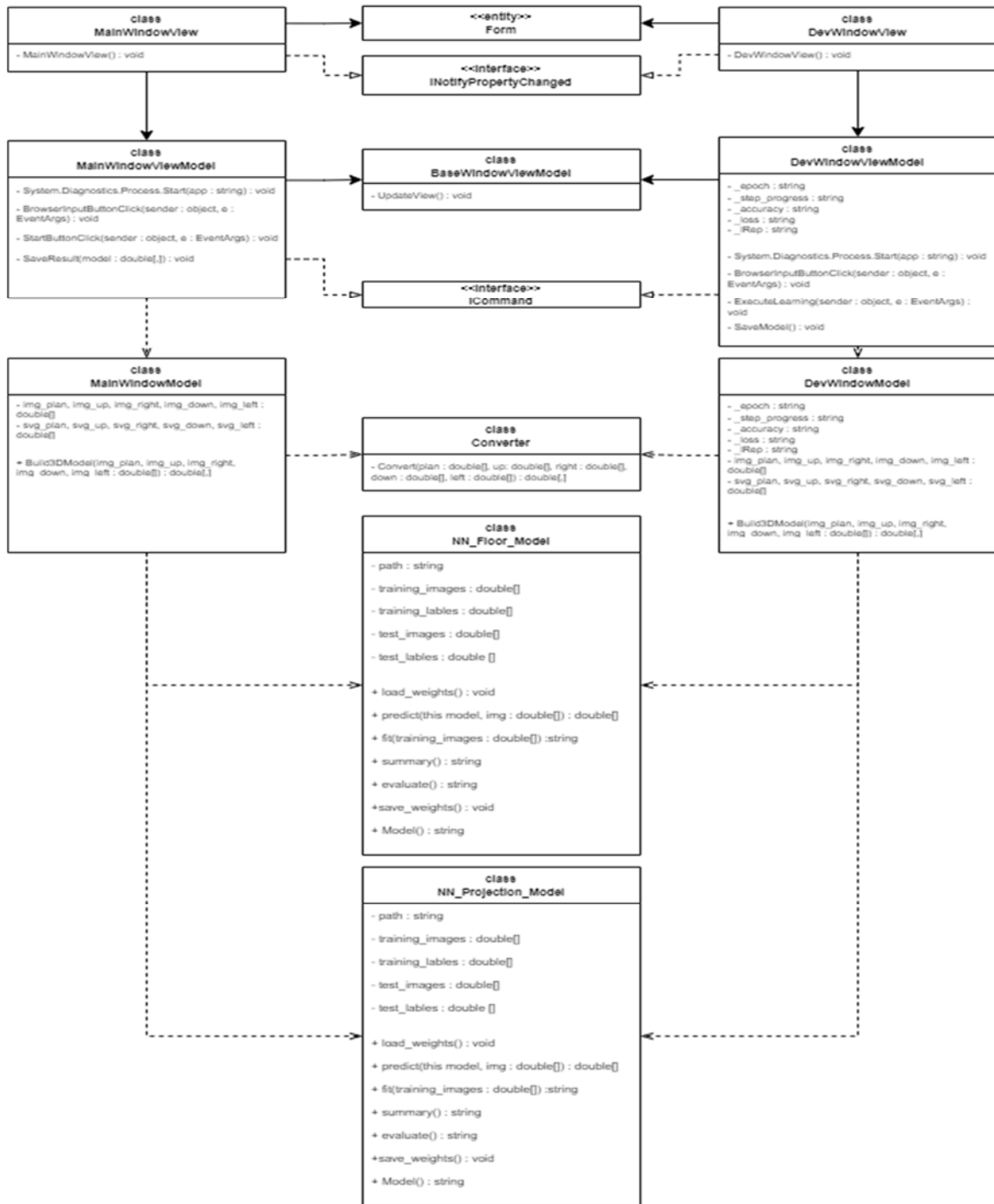


Рисунок 3.10. Діаграма класів додатку

Джерело: побудовано автором

Розглянемо детальніше кожен клас:

*MainWindowView*: Цей клас представляє головне вікно додатку. Він відповідає за відображення візуальних елементів, таких як кнопки, текстові поля та інші віджети. Він також слухає зміни властивостей моделі представлення через інтерфейс *INotifyProperty*, що дозволяє йому оновлювати відображення, коли змінюються дані.

*DevWindowView*: Цей клас представляє вікно розробника, яке може використовуватися для налаштування параметрів конвертації та перегляду внутрішніх даних. Він також слухає зміни властивостей моделі представлення через інтерфейс *INotifyProperty*.

*MainWindowViewModel*: Цей клас є посередником між *MainWindowView* та *MainWindowModel*. Він обробляє команди від представлення, такі як кліки на кнопках, та оновлює модель відповідно. Він також відповідає за обробку помилок та відображення відповідних повідомлень користувачеві.

*DevWindowViewModel*: Аналогічно *MainWindowViewModel*, цей клас є посередником між *DevWindowView* та *DevWindowModel*. Він може обробляти додаткові команди, специфічні для вікна розробника.

*MainWindowModel*: Цей клас відповідає за зберігання даних, що відображаються в *MainWindowView*, та взаємодію з зовнішніми ресурсами, такими як бази даних або веб-сервіси. Він також використовує класи *Converter*, *NN\_Floor\_Model* та *NN\_Projection\_Model* для обробки даних.

*DevWindowModel*: Цей клас відповідає за зберігання даних, що відображаються в *DevWindowView*, та може включати додаткову логіку для обробки даних, специфічних для вікна розробника.

Подібна архітектура дозволяє розділити логіку відображення, обробки даних та взаємодії з користувачем, що спрощує розробку та тестування додатку.

Для створення конкретної фізичної системи необхідно реалізувати всі елементи логічного опису в конкретних матеріальних елементах. Для опису



таких реальних елементів призначене фізичне подання моделі. У мові UML це означає сукупність зв'язаних елементів, включаючи програмне і апаратне забезпечення, а також персонал, які організовані для виконання спеціальних завдань. Для фізичного подання моделей систем використовуються діаграми реалізації, які включають дві окремі канонічні діаграми: діаграму компонентів і діаграму розгортання [91].

Діаграма компонентів дозволяє визначити архітектуру системи, що розробляється. Основними графічними елементами діаграми компонентів є компоненти, інтерфейси і залежності між ними [91].

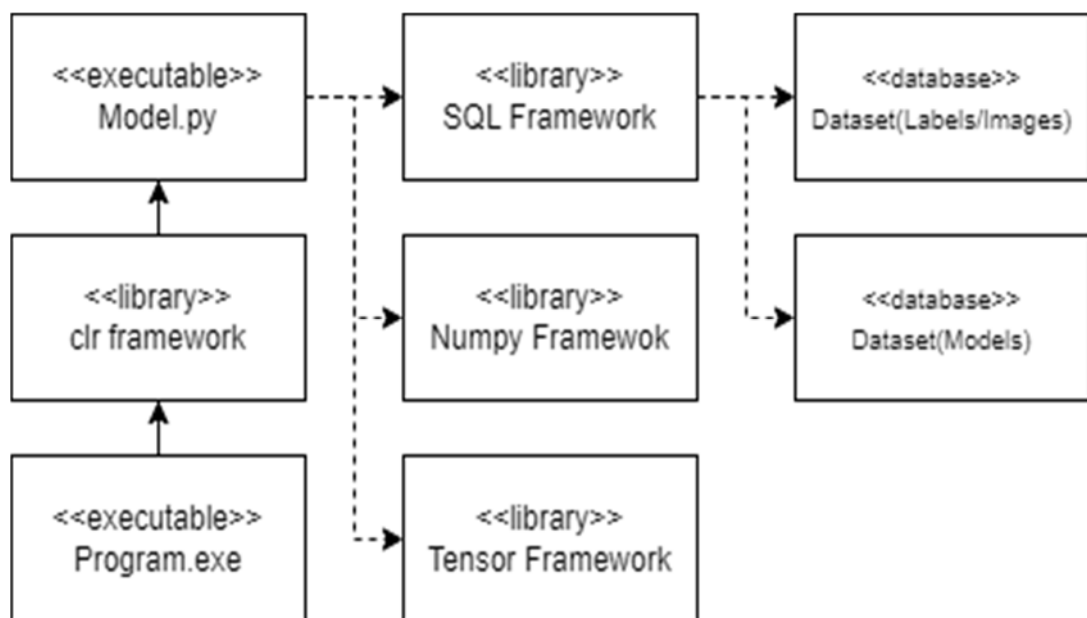


Рисунок 3.11. Діаграма компонентів додатку

*Джерело: побудовано автором*

Опишемо компоненти діаграми компонентів додатку інформаційної технології розпізнавання графічної інформації на основі нейронної мережі:

*Program.exe*: головний виконуваний файл програми, який слугує точкою входу для користувача. Він відповідає за ініціалізацію програми, взаємодію з користувачем та управління загальним потоком роботи програми. *Program.exe* також відповідає за запуск *Model.py* через CLR Framework.

*CLR Framework: Common Language Runtime (CLR)* - компонент .NET Framework від Microsoft, який дозволяє виконувати код, написаний на різних мовах програмування, в одному спільному середовищі виконання. В нашому випадку, CLR Framework дозволяє Program.exe запускати скрипт Python (Model.py).

*Model.py*: скрипт Python, який виконує основну роботу по конвертації 2D зображень в 3D моделі. Він використовує Tensor Framework для створення та навчання нейронних мереж, Numpy Framework для обробки числових даних, та SQL Framework для взаємодії з базами даних.

*Tensor Framework*: фреймворк для створення та навчання нейронних мереж. Він використовується Model.py для обробки 2D зображень та їх конвертації в 3D моделі.

*Numpy Framework*: бібліотека Python для обробки числових даних. Вона використовується Model.py для обробки даних, отриманих від Tensor Framework.

*SQL Framework*: бібліотека Python для взаємодії з SQL базами даних. Вона використовується Model.py для зберігання та отримання даних з баз даних Image and Label Dataset та Model Dataset.

*Image and Label Dataset, Model Dataset*: Ці дві бази даних використовуються Model.py для зберігання та отримання даних. Image and Label Dataset містить зображення та відповідні їм мітки, які використовуються для навчання нейронних мереж, а Model Dataset містить навчені моделі, які використовуються для перетворення 2D зображень в 3D моделі.

Розроблена архітектура програмного додатку інформаційної технології розпізнавання графічної інформації на основі нейронної мережі дозволяє досягти кращих результатів і робити розпізнавання більш точним і швидшим.

Дана інформаційна технологія може бути застосована до будь-якого графічного об'єкту, що має компонентну структуру та може бути представлений у вигляді графічної інформації (креслярської документації).

### ВИСНОВКИ ДО РОЗДІЛУ 3

Таким чином, в результаті досліджень, виконаних у третьому розділі роботи, були отримані такі результати:

- Розроблена концептуальна модель інформаційної технології розпізнавання графічної інформації на основі нейронної мережі, яка відображає всі етапи та процеси перетворення із 2D зображення в 3D модель. Дана модель відрізняється від існуючих подібних аналогів, саме відмінністю отриманої інформації, яка буде більш повною і більш широко розкриватиме можливість перенести точні деталі для 3D моделі.
- Спроектвана засобами CASE-технології ERwin інформаційна технологія розпізнавання графічної інформації на основі нейронної мережі з визначеними зовнішніми впливами на неї та взаємодії процесів всередині системи. Побудовані декомпозиції моделі у нотації IDEF0 та IDEF3 показують інформаційну технологію як в цілому, так і у вигляді простих проєкцій, які деталізовано описані. При чому кожен попередній етап декомпозиції безпосередньо впливає на всі наступні рівні декомпозиції та на прийняття рішень.
- Описана генеза інформаційної технології розпізнавання графічної інформації на основі нейронної мережі як комплекс методик та концепції, які застосовуються для розпізнавання графічної інформації та перетворення її із двовекторного простору у тривекторний.
- Показано, що основою інформаційної технології розпізнавання графічної інформації є нейронна мережа, основною задачею якої є розпізнавання та трансформація графічної інформації у тривимірний простір.
- Запропонована архітектура програмної реалізації концептуальної моделі інформаційної технології розпізнавання графічної інформації на основі нейронної мережі.

Основні результати розділу висвітлені у науково-дослідній роботі НДР № 0122U001549 «Модельовання інтелектуальних систем управління діяльністю підприємств», (довідка від 29.08.2023 №1453/24)

Основні результати розділу опубліковані в наукових працях автора:

Наукові праці, в яких опубліковані основні наукові результати дисертації:

[5], [6], [7], [9].

Наукові праці, які засвідчують апробацію матеріалів дисертації:

[10], [12], [13].

Наукові праці, які додатково відображають наукові результати дисертації

[14], [15], [16], [18], [19], [22].

## ВИСНОВКИ

У результаті проведеного дисертаційного дослідження на тему «Інформаційна технологія розпізнавання графічної інформації на основі нейронної мережі» розв'язано науково-прикладне завдання створення та впровадження моделі інформаційної технології розпізнавання графічної інформації на основі нейронної мережі.

1. Встановлено, що проблематика розпізнавання креслень та трансформації  $2D \rightarrow 3D$  є дуже актуальною в умовах розвитку різних інформаційних технологій та прикладного програмного забезпечення для систем автоматизованого проєктування, зокрема, орієнтованих на використання 3D моделей у ході вирішення прикладних технічних, технологічних, наукових та інших завдань. Однак, науково-технічне завдання, пов'язане з розпізнаванням креслень та трансформацією  $2D \rightarrow 3D$  на основі проєкційних зображень, залишається не до кінця вирішеним. Потребують розвитку як її алгоритмічна складова, так і програмно-практична частина, орієнтована на створення програмних продуктів.

Систематизовано та узагальнено методи, моделі, алгоритми та програмне забезпечення в задачах розпізнавання графічної інформації, що, дозволяють використовувати згорткові нейронні мережі в задачі розпізнавання креслень.

2. Досліджено потенціал застосування різних методів штучного інтелекту в задачі розпізнавання креслень та трансформації  $2D \rightarrow 3D$ . Науковий і практичний інтерес до даної тематики зумовлює велику кількість публікацій, пов'язаних із цим напрямком. Показано, що проблематика трансформації  $2D \rightarrow 3D$  може бути ефективно вирішена на основі застосування систем штучного інтелекту, зокрема нейронних мереж, що дозволить суттєво знизити трудомісткість вирішення окремих завдань як в алгоритмічному, так і в практичному аспектах.

3. Проаналізовано та узагальнено існуючі підходи до трансформації у вигляді етапів трансформації 2D зображень у 3D модель, які є узагальненням моделі автоматизованої підготовки виробництва для задачі розпізнавання креслень, і сприяють подальшій автоматизації проектно-конструкторських та проектно-технологічних робіт.

4. Розроблено модель класифікатора графічних зображень на основі покриттів класів та елементарних класифікаторів примітивів для підвищення ефективності навчання CNN.

Показано, що конструювання множини елементарних класифікаторів для креслярських примітивів зводиться до знаходження допустимих та максимальних кон'юнкцій для функції, що описує аналізований клас графічних об'єктів.

5. Розроблено архітектуру згорткової нейронної мережі з модулем елементарних класифікаторів графічних примітивів для вирішення задачі розпізнавання креслень та трансформації  $2D \rightarrow 3D$ .

Показано, що парадигмою побудови бази примітивів на навчання CNN чи логічних процедур навчання є пошук інформативних фрагментів описів відповідних графічних примітивів. Ці фрагменти при створенні конкретних проектних рішень для систем розпізнавання креслень, дозволять однозначно робити висновок про наявність (або відсутність) в базі системи розпізнавання певного графічного примітиву в рамках класу.

Викладено основні принципи конструювання логічних процедур навчання з використанням апарату логічних функцій, що дозволить на практиці створювати ефективні програмні рішення CNN у задачах розпізнавання креслень та трансформації  $2D \rightarrow 3D$ .

6. Визначена генеза інформаційної технології розпізнавання графічної інформації на основі нейронної мережі як комплекс методик та концепції, які застосовуються для розпізнавання графічної інформації та перетворення її із двовекторного простору у тривекторний.

Показано, що базовою структурою інформаційної технології розпізнавання графічної інформації є нейронна мережа, основною задачею якої є розпізнавання та трансформація графічної інформації у тривимірний простір.

Застосовано нейромережевий підхід для розпізнавання та трансформації графічної інформації у тривимірний простір за допомогою згорткової нейронної мережі з модулем елементарних класифікаторів графічних примітивів для вирішення задачі розпізнавання креслень та трансформації графічної інформації із двовимірного простору у тривимірний.

7. Розроблена концептуальна модель інформаційної технології розпізнавання графічної інформації на основі нейронної мережі, яка відображає всі етапи та процеси перетворення 2D зображення в 3D модель. Дана модель представляє отриману інформацію у більш повному форматі, що дозволяє розкрити потенційні можливості перенесення точних деталей зображення у 3D модель.

Спроектовано засобами CASE-технології ERwin інформаційну технологію розпізнавання графічної інформації на основі нейронної мережі з визначеними зовнішніми впливами на неї та взаємодією процесів всередині системи. Побудовані декомпозиції моделі у нотації IDEF0 та IDEF3 представляють інформаційну технологію як в цілому, так і у вигляді простих деталізованих проєкцій. У такому представленні кожен попередній етап декомпозиції безпосередньо впливає на всі наступні рівні декомпозиції та на процес прийняття рішень.

Запропонована архітектура програмної реалізації концептуальної моделі інформаційної технології розпізнавання графічної інформації на основі нейронної мережі.

Розроблена інформаційна технологія розпізнавання може бути застосована до будь-якого графічного об'єкту, що має компонентну



структуру і представлений у вигляді графічної інформації (креслярської документації).

Практична значущість результатів досліджень полягає в тому, що розроблена, спроектована та змодельована інформаційна технологія розпізнавання графічної інформації на основі нейронної мережі, яка може бути застосована до будь-якого графічного об'єкту, що має компонентну структуру та може бути представлений у вигляді графічної інформації (креслярської документації).

Реалізація даної технології була впроваджена у виробничу діяльність підприємств: ТОВ «ФЬЮЧЕ КОНСАЛТИНГ», ТОВ «ІНТЕРНЕТ ІНВЕСТИЦІЙНА ГРУПА».

Подальші дослідження передбачають розвиток удосконалення алгоритмів нейромережі для підвищення точності та повноти розпізнавання графічної інформації та удосконалення процесу трансформації з двовимірногоу тривимірний простір, що виражено кількістю обробленої графічної інформації за визначену одиницю часу.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Котлик, С., Соколова, О., & Корнієнко, Ю. ОНАХТ (Україна) Огляд застосовування програмного забезпечення для 3D моделювання. XII Міжнародна науково-практична конференція інформаційні технології і автоматизація – 2019. 58-60.
2. Пузь Д.О. Застосування 3D моделювання в сферах людської діяльності [Електронний ресурс] / Д.О. Пузь, С.В. Сомов // Новітні інформаційні системи та технології - Modern information system and technologies. – Полтава : ПолтНТУ, 2018. – Вип. 9.
3. Sutherland, I. E. (1964). Sketchpad a man-machine graphical communication system. *Simulation*, 2(5), R-3.
4. De Villiers, M. (1999). The role and function of proof with Sketchpad. *Rethinking proof with Sketchpad*, 3-10.
5. Okino, T., Murata, H., Taima, K., Iinuma, T., & Oketani, K. (1996, April). New television with 2D/3D image conversion technologies. In *Stereoscopic Displays and Virtual Reality Systems III* (Vol. 2653, pp. 96-103). SPIE.
6. Sisi, L., Fei, W., & Wei, L. (2010, November). The overview of 2D to 3D conversion system. In *2010 IEEE 11th International Conference on Computer-Aided Industrial Design & Conceptual Design 1* (Vol. 2, pp. 1388-1392). IEEE.
7. Fan, Y. C., Chiu, Y. C., & Chang, L. C. (2022, June). 2D/3D Image Converter Based on Overlapping Line. In *2022 IEEE International Conference on Imaging Systems and Techniques (IST)* (pp. 1-4). IEEE.
8. Tarjan, R. An Efficient Planarity Algorithm / R. Tarjan // Computer Science Department. – Report № CS – 244 – 71: Stanford University, November, 1971.

9. Cobb, E.C. On the Extraction of Solid Geometry from a Wire Frame Geometric Data Base / E.C. Cobb // M.S. Thesis: University of Wisconsin-Madison. – 1978.
10. Lequette, R. Automatic construction of curvilinear solids from wire frame views / R. Lequette // France. – 1988. – Vol.20, №4. – P.171–178.
11. Requicha, Aristides AG, and Herbert B. Voelcker. "Constructive solid geometry." (1977).
12. Breitenberger, M., et al. "Analysis in computer aided design: Nonlinear isogeometric B-Rep analysis of shell structures. "Computer Methods in Applied Mechanics and Engineering, 284 (2015): 401-457.
13. Zhu, H., & Menq, C. H. (2002). B-rep model simplification by automatic fillet/round suppressing for efficient automatic feature recognition. Computer-Aided Design, 34(2), 109-123.
14. Idesawa, M. A Automatic Input of Line Drawing and Generation of Solid Figure from Three-View Data / M.A. Idesawa, T. Soma, E. Goto, S. Shibata // Proceed-ings of the International Joint Comput-er Symposium, 1975. – P. 304 – 311.
15. Idesawa, M. A System to Generate a Solid Figure from a Three View / M.A. Idesawa, Bull. JSME 16. – February, 1973. – P. 216 – 225.
16. Lafue, G. Recognition of Three Dimensional Objects from Views / G. Lafue // Computer Graphics. – 1976. – Vol. 10, № 2.
17. Preiss, K. Constructing the 3-D Representation of a Plane-Faced Object from a Digitized Engineering Drawings / K. Preiss // Fifth International Conference and Exhibition on Computer in Engineering and Building Design, 1980.
18. Preiss, K. Constructing the Solid representation from engineering projections / K. Preiss // Computer & Graphics. – 1984. – Vol. 8, № 4. – P. 381 – 389.
19. Haralick, R.M. Understanding engineering drawings / R.M. Haralick, D. Queeney // Comp. Graphics and Image Processing. – 1982. – Vol.20, № 3. – P. 244 – 258.

20. Markowsky G. Fleshing out projections / G. Markowsky, M.A. Wesley // IBM J. Res& Develop. November, 1981. – Vol. 25, № 6. – P. 934 – 954.
21. Markowsky, G. Fleshing out wire frames / G. Markowsky, M.A. Wesley // IBM J. Res.& Develop. – sept. 1980. – Vol. 24, № 5. – P. 582 – 587.
22. Markowsky, G. Generation of solid models from two-dimensional and three-dimensional data / G. Markowsky, M.A. Wesley // in Pickett, MS and Boyse, J M (eds). Solid modelling by computer: from theory to application: Plenum, 1986. – P. 23 – 51.
23. Sakurai, H. Solid Model Input Through Orthographic Views / H. Sakurai, D.C. Gossard // Computer Graphics. – 1983. – Vol.17, №3. – P. 243 – 252.
24. Wanga Z., Tsumuraa K., Saitoa Y. Recognition of hand-written mechanical drawing by multilevel neural network (MLNN) // Seventh International Conference on Production/Precision Engineering, 4th International Conference on High Technology. Chiba Japan: 1994, pp. 1-6.
25. Elyan E., Jamieson L., Ali-Gombe A. Deep learning for symbols detection and classification in engineering drawings // Neural Networks. September 2020, no. 129, pp. 91-102.
26. Кубик, Олександр Олексійович, Олександр Вікторович Мазурець, and Сергій Станіславович Ковальчук. "Декомпозитивне розпізнавання символічної інформації з креслень із використанням технологій штучного інтелекту." Математичне та комп'ютерне моделювання. Серія: Технічні науки (2008): 109-119.
27. Zehtaban L., Elazhary O., Roller D. A framework for similarity recognition of CAD models // Journal of Computational Design and Engineering, 2016, no. 3, pp. 274-285.
28. Wang Z., Latif M. Reconstruction of 3D Solid Models Using Fuzzy Logic Recognition//Proceedings of the World Congress on Engineering. 2007. V. 1. pp. 37-42.

29. Governi L., Furferi R., Palai M., Volpe Y. 3D Geometry Reconstruction from Orthographic Views: a Method Based on 3D Image Processing and Data Fitting // *Computers in Industry*. 2013. № 64. P. 1290-1300.
30. Nagendra, I.V. 3D Objects from 2D orthographic views – A Survey / I.V. Nagendra, U.G. Gujar // *Computer & Graphics*. – 1988. – Vol.12, №1. – P. 111 – 114.
31. Governi L., Furferi R., Palai M., Volpe Y. 3D Geometry Reconstruction from Orthographic Views: a Method Based on 3D Image Processing and Data Fitting // *Computers in Industry*. 2013. № 64. P. 1290-1300.
32. Kargas, A. Interpretation of engineering drawings as solid models / A. Kargas, P. Cooley, T.H.E. Richards // *Computer-Aided Engineering Journal*. – april 1988. – P. 67 – 78.
33. EUCLID. Technical Report. MATRA Datavision, France, 1988. –200 p.
34. Omura, G. (2012). *Mastering AutoCAD 2013 and AutoCAD LT 2013*. John Wiley & Sons.
35. Avetisyan, A., Dahnert, M., Dai, A., Savva, M., Chang, A. X., & Nießner, M. (2019). Scan2cad: Learning cad model alignment in rgb-d scans. In *Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition* (pp. 2614-2623).
36. Yuan, S., Chan, H. S., Filipek, S., & Vogel, H. (2016). PyMOL and Inkscape bridge the data and the data visualization. *Structure*, 24(12), 2041-2042.
37. Sutherland, I.E. SCETCHPAD: A Man-Machine Graphical Communication System / I.E. Sutherland. – Proc. SJCC 23, 1963. – P. 329.
38. Thornton, R.W. Interactive Modelling in Three Dimensions through Two-Dimensional Windows / R.W. Thornton // *Third International Conference and Exhibition on Computer in Engineering and Building Design*, 1978.
39. Woo, T.C. Recognition of Three Dimensional designs from orthographic projections / T.C. Woo, J.M. Hammer // *Proc. 9th CIRP Conference: Cranfield Institute of Technology, Cranfield, England, 1977*. – P.247 – 255.

40. Dengel, A., & Spitz, A. L. (Eds.). (1995). Document Analysis Systems- Proceedings Of The International Association For Pattern Recognition Workshop (Vol. 14). World Scientific.

41. B. Pasternak. Processing imprecise and structural distorted line drawings by an adaptable drawing interpretation kernel. In A. Dengel and L. Spitz, editors, Proc. of IAPR Workshop on Document Analysis Systems, pages 349–366, 1994.

42. Messmer, B. T., & Bunke, H. (1995, August). Automatic learning and recognition of graphical symbols in engineering drawings. In International Workshop on Graphics Recognition (pp. 123-134). Springer, Berlin, Heidelberg.

43. Castellano, G., & Vessio, G. (2021). Deep learning approaches to pattern extraction and recognition in paintings and drawings: An overview. *Neural Computing and Applications*, 33(19), 12263-12282.

44. Попроцька Д. І. Інформаційна система розпізнавання креслень / Д. І. Попроцька ; наук. кер. О. А. Шпинковський // Стан, досягнення і перспективи інформаційних систем і технологій : матеріали XX Всеукр. наук.-техн. конф. молодих вчених, аспірантів та студентів, Одеса, 21-22 квіт. 2020 р. / Одес. нац. акад. харч. технологій, ННІ комп'ютер. систем і технологій «Індустрія 4.0» ім. П. М. Платонова, Фак. комп'ютер. інженерії, програмування та кіберзахисту ; орг. ком.: Б. В. Єгоров (голова) та ін. – Одеса : ОНАХТ, 2020. – Ч. I. – С. 166–168. – Бібліогр.: 4 назв

45. Ковальчук, С. С., & Мазурець, О. В. (2008). Аналіз ефективності використання технології багат шарових нейросхем для вирішення складних технологічних задач.

46. Seo, Jiho, Hyejin Park, and Seungyeon Choo. "Inference of drawing elements and space usage on architectural drawings using semantic segmentation." *Applied Sciences* 10.20 (2020): 7347.

47. Shaaban, A. M., Salem, N. M., & Al-atabany, W. I. (2020). A semantic-based scene segmentation using convolutional neural networks. *AEU-International Journal of Electronics and Communications*, 125, 153364.

48. Fonseca, M. J., & Jorge, J. A. (2000, May). Using fuzzy logic to recognize geometric shapes interactively. In Ninth IEEE International Conference on Fuzzy Systems. FUZZ-IEEE 2000 (Cat. No. 00CH37063) (Vol. 1, pp. 291-296). IEEE.
49. Adamik, M., Goga, J., Pavlovicova, J., Babinec, A., & Sekaj, I. (2022). Fast robotic pencil drawing based on image evolution by means of genetic algorithm. *Robotics and Autonomous Systems*, 148, 103912.
50. Wang, Z., & Latif, M. (2007, July). Reconstruction of 3D Solid Models Using Fuzzy Logic Recognition. In *World Congress on Engineering* (pp. 37-42).
51. Jin, Bingwen, Songhua Xu, and Weidong Geng. "Learning to sketch human facial portraits using personal styles by case-based reasoning." *Multimedia Tools and Applications* 77.5 (2018): 5417-5441.
52. Wu J., Zhang C., Xue T., Freeman W. T., Tenenbaum J. B. Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling, In *NIPS*, pages 82–90, 2016.
53. Häne, C., Tulsiani, S., & Malik, J. (2017). Hierarchical Surface Prediction for 3D Object Reconstruction. *2017 International Conference on 3D Vision (3DV)*, 412-420. DOI:10.1109/3DV.2017.00054
54. Khorolska, K., Bebeshko, B., Desiatko, A., Lazorenko, V. 3D Models Classification with Use of Convolution Neural Network *CEUR Workshop Proceedings*, 2021, 3179, pp. 25–34 [http://ceur-ws.org/Vol-3179/Paper\\_3.pdf](http://ceur-ws.org/Vol-3179/Paper_3.pdf)
55. Wu, Z., Song, S., Khosla, A., Yu, F., Zhang, L., Tang, X., & Xiao, J. (2015). 3D ShapeNets: A deep representation for volumetric shapes. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1912-1920. DOI:10.1109/CVPR.2015.7298801
56. Zhirong Wu, Shuran Song, Aditya Khosla, Fisher Yu, Linguang Zhang, Xiaoou Tang, and Jianxiong Xiao. 3d shapenets: A deep representation for volumetric shapes. In *CVPR*, pages 19121920, 2015. <https://arxiv.org/pdf/1406.5670.pdf>

57. Feng, Y., Zhang, Z., Zhao, X., Ji, R., & Gao, Y. (2018). GVCNN: Group-View Convolutional Neural Networks for 3D Shape Recognition. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 264-272. DOI:10.1109/CVPR.2018.00035
58. B. K. P. Horn, "Extended Gaussian images," in Proceedings of the IEEE, vol. 72, no. 12, pp. 1 1686, Dec. 1984, doi: 10.1109/PROC.1984.13073.
59. Bundzel M., Hashimoto S. Object identification in dynamic images based on the memory prediction theory of brain function // Journal of Intelligent Learning Systems and Applications. – 2010. – V. 2. – № 4. – P. 212–220.
60. Rowley H.A., Baluja S., Kanade T. Neural network based face detection // Pattern anal. mach. intell. – 1998. – V. 20. – P. 23–38.
61. Feraud R., Bernier O., Viallet J., Collobert M. A fast and accurate face detector based on neural networks // IEEE Transactions on pattern analysis and machine intelligence. – 2001. – V. 23. – № 23. – P. 42–53.
62. Le Cun Y., Kavukcuoglu K., Farabet C. Convolutional Networks and Applications in Vision // International Symposium on Circuits and Systems (ISCAS'10). – Paris, France: IEEE Computer Society, 2010. – P. 253–256.
63. LeCun Y. and Y. Bengio. "Convolutional networks for images, speech and time series", in M.A. Arbib, ed., The Hand book of Brain Theory and Neural Networks, Cambridge, MA: MIT Press, 1995.
64. Karen Simonyan, Andrew Zisserman, 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
65. Sergey Ioffe, Christian Szegedy, 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167.
66. Deep Residual Learning for Image Recognition / Kaiming He [et al.]. – In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 770–778, 2016.



67. Generalizing pooling functions in convolutional neural networks: Mixed, gated, and tree / Chen-Yu Lee [et al.]. – In *Artificial Intelligence and Statistics*, pages 464–472, 2016.
68. K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. In *ICLR*, 2015.
69. Bergevin, R., & Levine, M. D. (1993). Generic object recognition: Building and matching coarse descriptions from line drawings. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(1), 19-36.
70. Yin, R., Monson, E., Honig, E., Daubechies, I., & Maggioni, M. (2016, March). Object recognition in art drawings: Transfer of a neural network. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 2299-2303). IEEE.
71. Hiraoka, K., Suzuki, K., Hirayama, K., & Mori, E. (2009). Visual agnosia for line drawings and silhouettes without apparent impairment of real-object recognition: a case report. *Behavioural neurology*, 21(3-4), 187-192.
72. Heuer, S. (2016). The influence of image characteristics on image recognition: A comparison of photographs and line drawings. *Aphasiology*, 30(8), 943-961.
73. Gupta, S., Kumar, M., & Garg, A. (2019). Improved object recognition results using SIFT and ORB feature detector. *Multimedia Tools and Applications*, 78(23), 34157-34171.
74. Gadat, S. (2005, July). Extraction of attributes for visual object recognition and dna microarray analysis. In *IEEE/SP 13th Workshop on Statistical Signal Processing, 2005* (pp. 1370-1375). IEEE.
75. Froba, B., & Ernst, A. (2004, May). Face detection with the modified census transform. In *Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. Proceedings.* (pp. 91-96). IEEE.
76. LeCun, Y., Haffner, P., Bottou, L., & Bengio, Y. (1999). Object recognition with gradient-based learning. In *Shape, contour and grouping in computer vision* (pp. 319-345). Springer, Berlin, Heidelberg.

77. Бекетова, Р., Ахметов, Б., Корченко, О., & Лахно, В. (2016). Розробка моделі інтелектуального розпізнавання аномалій та кібератак з використанням логічних процедур, що базуються на покриттях матриць ознак. *Безпека інформації*, (22, № 3), 242-254.
78. Chen, Z. M., Wei, X. S., Wang, P., & Guo, Y. (2019). Multi-label image recognition with graph convolutional networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 5177-5186).
79. Yu, Z., Li, T., Luo, G., Fujita, H., Yu, N., & Pan, Y. (2018). Convolutional networks with cross-layer neurons for image recognition. *Information Sciences*, 433, 241-254.
80. Bhuvaneshwari, C., & Manjunathan, A. (2020). Advanced gesture recognition system using long-term recurrent convolution network. *Materials Today: Proceedings*, 21, 731-733.
81. Liu, J. J., Hou, Q., Cheng, M. M., Wang, C., & Feng, J. (2020). Improving convolutional networks with self-calibrated convolutions. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 10096-10105).
82. Imambi, S., Prakash, K. B., & Kanagachidambaresan, G. R. (2021). PyTorch. In *Programming with TensorFlow* (pp. 87-104). Springer, Cham.
83. Vasilev, I., Slater, D., Spacagna, G., Roelants, P., & Zocca, V. (2019). *Python Deep Learning: Exploring deep learning techniques and neural network architectures with Pytorch, Keras, and TensorFlow*. Packt Publishing Ltd.
84. Agrawal, A., Modi, A., Passos, A., Lavoie, A., Agarwal, A., Shankar, A., ... & Cai, S. (2019). TensorFlow Eager: A multi-stage, Python-embedded DSL for machine learning. *Proceedings of Machine Learning and Systems*, 1, 178-189.
85. Brownlee, J. (2016). *Deep learning with Python: develop deep learning models on Theano and TensorFlow using Keras*. Machine Learning Mastery.
86. Marr D., Hildreth E. Theory of edge detection. *Proc. R. Soc. (London)*, 1980, B 207, pp. 187–217.

87. Технології 3D друку. URL: <https://www.techpats.com/3d-printingtechnologies-overview/>

88. Ващенко А.С. Огляд методів моделювання 3D об'єктів. Інформаційні технології та системи управління в економіці, техніці та екології Інформаційні технології та системи управління в економіці, техніці та екології: тези доп. міжнар. наук.-практ. конф., м. Харків, 20-21 лют. 2020 р. Харків, 2020. С. 35.

89. Virtual Evolution: An Alternate History of Cyberspace, 1988-1997. URL: <http://search.proquest.com/openview/551852c98713d7d104b584dcec403384/1?pqorigsite=gscholar&cbl=51922&diss=y>

90. Zhiqin Chen and Hao Zhang. Learning implicit fields for generative shape modeling. arXiv preprint arXiv:1812.02822, 2018.

91. В.М.Дубовой, С.М.Москвіна, О.Д.Никитенко. Моделювання процесів і систем керування. Вінниця – ВНТУ, 2009

92. Закон України «Про інформацію». URL: <https://zakon.rada.gov.ua/laws/show/2657-12#Text>

93. Закон України «Про захист інформації в інформаційно-комунікаційних системах». URL: <https://zakon.rada.gov.ua/laws/show/80/94-%D0%B2%D1%80#Text>

94. Закон України «Про Національну програму інформатизації». URL: <https://zakon.rada.gov.ua/laws/show/74/98-%D0%B2%D1%80#Text>

95. Закон України «Про електронні комунікації». URL: <https://zakon.rada.gov.ua/laws/show/1089-20#Text>

## ДОДАТКИ

Таблиця А1

## Складові компоненти «Інформаційна технологія розпізнавання графічної інформації»

Назва (Name)	Характеристика (Description)
Підготовка даних (Data preparation)	Розподілення вхідної інформації як первинний етап в навчанні нейронної мережі: перетворення технічного креслення у зображення двовимірного простору. Процес навчання (з подальшим тестуванням) нейронної мережі будується на основі сприйняття зображення на рівні пікселя і призначення кожному пікселю класу об'єктів відповідно до шаблону, який нейронна мережа створює шляхом ітерації набору даних. Для виділення архітектурних елементів використовується CNN (Convolutional Neural Networks).
Оброблення інформації (Information extraction)	Розпізнавання на основі алгоритмів CNN введеного користувачем технічного креслення (графічної інформації) та перетворення його з двовекторного формату у тривекторний для подальшої побудови 3D моделі.
Побудова 3D моделі (3D model construction)	Моделювання 3D образу розпізнаної графічної інформації на основі отриманих даних від нейромережі.

Таблиця А2

## Характеристика стрілок до діаграми «Інформаційна технологія розпізнавання графічної інформації»

Назва (Name)		Характеристика (Description)	Тип (Type)	Складові моделі (Components)
Креслення		Фронтальний вигляд об'єкта, бокове зображення, зображення ззаду, план зверху.	Input	Data preparation
		Цифровий формат зображень зовнішніх частин об'єкту, що потребує розпізнавання та проєктування у 3D форматі	Input	Information extraction
Нормативно-правова база	Закон України «Про Національну програму інформатизації»	Національна програма інформатизації становить комплекс взаємопов'язаних окремих завдань (проєктів) інформатизації, спрямованих на реалізацію державної політики та пріоритетних напрямів створення сучасної інформаційної інфраструктури України за рахунок концентрації та раціонального використання фінансових, матеріально-технічних та інших ресурсів, виробничого і науково-технічного потенціалу держави, а також координації діяльності державних органів, органів місцевого самоврядування, підприємств, установ, організацій усіх форм власності і громадян у сфері інформатизації [94].	Control	3D model construction
	Закон України «Про електронні комунікації»	Цей Закон регулює відносини у сфері захисту інформації в інформаційних, телекомунікаційних та інформаційно-телекомунікаційних системах (далі - система) [93].	Control	Information extraction

	Закон України «Про інформацію»	Цей Закон регулює відносини щодо створення, збирання, одержання, зберігання, використання, поширення, охорони, захисту інформації [92].	Control	Data preparation
Генеза розпізнавання графічної інформації	Методика підготовки даних для навчання нейронної мережі	Комплекс правил та рекомендацій щодо підготовки первинних даних для навчання нейронної мережі.	Control	Data preparation
	Методика навчання та тестування нейронної мережі	Комплекс програмних алгоритмів для формування бази даних та бази знань з метою розробки механізмів розпізнавання графічної інформації для переведення інформації у двовекторний простір. Тестування нейронної мережі відбувається після попереднього процесу навчання з метою виявлення похибок та подальшим їх коригуванням.	Control	Information extraction
	Концепція побудови 3D моделі	Система алгоритмів на основі яких нейронна мережа перетворює інформацію із двовекторного простору у тривекторний	Control	3D model construction
Програмно-апаратний комплекс	ПК	CPU - Intel Core i7-12700K GPU - GeForce RTX 3090 super RAM - Corsair Vengeance LPX 32GB (2 X 16GB) 3600MHz DDR4 C18 Storage - Samsung 980 500GB NVMe Gen 3 SSD Motherboard - Gigabyte Z690 AORUS MASTER	Mechanism	Data preparation
			Mechanism	Information extraction
			Mechanism	3D model construction
	PostgreSQL	Вибір обумовлений тим, що PostgreSQL є найліпшою продакшен базою даних з лінійки топ сегменту безкоштовних баз даних. Вигідною особливістю постгре є вдала, можливо найкраща, робота з бінарними даними. Вона чудово підходить для збереження зображень оперативного рівня. Отримання їх з бази значно швидше ніж читання з файлу, навіть за використання сучасних NVMe Gen 3 SS.	Mechanism	Data preparation
	GoLang	Мова програмування вибрана з огляду на ряд факторів - 1) Відносно низький рівень входження в мову, що забезпечує максимальну “сумісність” з програмістами різного рівня кваліфікації 2) Має сучасний синтаксис та потужність низькорівневої мови програмування 3) Компілюється в машинний код, відповідно є максимально “нативною” та кросплатформенною за потреби. IDE - GoLand Вибір GoLand IDE обумовлений рядом фактів 1) Безкоштовна повна ліцензія для наукових цілей 2) Потужний інструментарій “з коробки” - встановив і використовуєш, система позбавляє користувача зайвих турбот.	Mechanism	Information extraction
Mechanism			3D model construction	

		3) Великі можливості в базі Мінусом вибраної IDE є вибагливість до наявних ресурсів ПК.		
Human resources	Database manager	Технічний працівник, який маркує первинні данні технічного креслення (графічної інформації), необхідний на етапі первинного навчання нейронної мережі	Mechanism	Data preparation
	Developer	Розробник нейронної мережі	Mechanism	Information extraction
	Customer	Кінцевий користувач технології розпізнавання графічної інформації	Mechanism Mechanism	Information extraction 3D model construction
Sorted DataSets		Відсортовані бази знань, які є первинними даними для навчання нейронної мережі	Output Input	Data preparation Information extraction
Vectorized grid image		Інтегрована 2D зображення з координатною трьохвимірною сіткою	Output Input	Information extraction 3D model construction
3D model		Результат оброблення та перетворення інформації із двовекторного простору у тривекторний нейронною мережею	Output	3D model construction

Таблиця А3

## Складові компоненти Data preparation

Назва (Name)	Характеристика (Description)
Surface blueprint	Розподілені двовекторні креслення: фронтальний вигляд об'єкта, бокове зображення, зображення ззаду, план зверху
Side view	Інтегровані 2D зображення: фронтальний вигляд об'єкта, бокове зображення, зображення ззаду

Таблиця А4

## Характеристика стрілок до діаграми «Data preparation»

Назва (Name)		Характеристика (Description)	Тип (Type)	Складові моделі (Components)
Креслення	Floor blueprint	Фронтальний вигляд об'єкта, бокове зображення, зображення ззаду, план зверху	Input	Surface blueprint
	Elevation	2D зображення	Input	Side view
Educated DataSets		Результат оброблення та перетворення інформації із двовекторного простору у тривекторний нейронною мережею: фронтальний вигляд об'єкта, бокове зображення, зображення ззаду, план зверху	Output	Surface blueprint
		Результат оброблення та перетворення 2D зображень із двовекторного простору у тривекторний нейронною мережею	Output	Side view
Закон України «Про інформацію»		Цей Закон регулює відносини у сфері захисту інформації в інформаційних, телекомунікаційних та інформаційно-	Control	Surface blueprint

		телекомунікаційних системах (далі - система) [93].		
		Цей Закон регулює відносини щодо створення, збирання, одержання, зберігання, використання, поширення, охорони, захисту інформації [92].	Control	Side view
Методика підготовки даних для навчання нейронної мережі	Алгоритм підготовки даних для навчання нейронної мережі	Комплекс програмних алгоритмів для формування бази даних та бази знань з метою розробки механізмів розпізнавання графічної інформації для переведення інформації у двовекторний простір. Тестування нейронної мережі відбувається після попереднього процесу навчання з метою виявлення похибок та подальшим їх коригуванням.	Control	Surface blueprint
			Control	Side view
ПК		CPU - Intel Core i7-12700K GPU - GeForce RTX 3090 super RAM - Corsair Vengeance LPX 32GB (2 X 16GB) 3600MHz DDR4 C18 Storage - Samsung 980 500GB NVMe Gen 3 SSD Motherboard - Gigabyte Z690 AORUS MASTER	Mechanism	Surface blueprint
			Mechanism	Side view
PostgresSQL		Найліпша продакшен базою даних з лінійки топ сегменту безкоштовних баз даних. Вигідною особливістю постгре є вдала, можливо найкраща, робота з бінарними даними. Вона чудово підходить для збереження зображень оперативного рівня. Отримання їх з бази значно швидше ніж читання з файлу, навіть за використання сучасних NVMe Gen 3 SSD	Mechanism	Surface blueprint
			Mechanism	Side view
Database/DataSet Manager		Технічний працівник, який маркує первинні данні технічного креслення (графічної інформації), необхідний на етапі первинного навчання нейронної мережі	Mechanism	Surface blueprint
			Mechanism	Side view



Довідка про впровадження в освітній процес ДТЕУ окремих положень, висновків та пропозицій, що містяться в дисертаційному дослідженні



МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
ДЕРЖАВНИЙ ТОРГОВЕЛЬНО-ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ

вул. Кіото, 19, м. Київ, 02156, тел. +380 (44) 531 47 41, e-mail: knute@knute.edu.ua, код ЄДРПОУ 44470624

30.08.2023 № 1458/22

На № \_\_\_\_\_

**ДОВІДКА**

Видана Хорольській Карині Вікторівні, здобувачу вищої освіти третього освітнього рівня «доктор філософії» (наукова спеціальність 122 «Комп'ютерні науки») кафедри інженерії програмного забезпечення та кібербезпеки Державного торговельно-економічного університету, про те, що окремі положення, висновки та пропозиції, що містяться в дисертаційному дослідженні на тему «Інформаційна технологія розпізнавання графічної інформації на основі нейронної мережі», застосовуються в освітньому процесі ДТЕУ при вкладанні дисциплін, що закріплені за кафедрою інженерії програмного забезпечення та кібербезпеки.

Хорольською Кариною Вікторівною розроблено (у співавторстві) та впроваджено в освітній процес:

1. «Експертні системи». Програма (КНТЕУ-2020). Для здобувачів вищої освіти ОС «бакалавр» спеціальностей 121 «Інженерія програмного забезпечення» та 125 «Кібербезпека».
2. «VR і 3D моделювання». Програма (КНТЕУ-2020). Для здобувачів вищої освіти ОС «магістр» спеціальності 121 «Інженерія програмного забезпечення».
3. «WEB-дизайн та WEB-програмування». Програма (КНТЕУ-2020). Для здобувачів вищої освіти ОС «бакалавр» спеціальності 121 «Інженерія програмного забезпечення».
4. «Управління проектами інформатизації». Програма (КНТЕУ-2020). Для здобувачів вищої освіти ОС «бакалавр» спеціальностей 121 «Інженерія програмного забезпечення» та 125 «Кібербезпека».
5. «WEB-дизайн та WEB-програмування». Робоча програма (КНТЕУ-2021). Для здобувачів вищої освіти ОС «бакалавр» спеціальності 242 «Туризм».
6. «Архітектура та проектування програмного забезпечення». Програма (КНТЕУ-2020). Для здобувачів вищої освіти ОС «бакалавр» спеціальностей 121 «Інженерія програмного забезпечення» та 125 «Кібербезпека».

7. «Архітектура та проектування програмного забезпечення». Робоча програма (ДТЕУ-2022). Для здобувачів вищої освіти ОС «бакалавр» спеціальності 121 «Інженерія програмного забезпечення».
8. «Архітектура та проектування програмного забезпечення». Робоча програма (ДТЕУ-2022). Для здобувачів вищої освіти ОС «бакалавр» спеціальності 125 «Кібербезпека».
9. «Хмарні та GRID-технології». Програма (КНТЕУ-2021). Для здобувачів вищої освіти ОС «бакалавр» спеціальності 124 «Системний аналіз», спеціальності 051 «Економіка», спеціальності 076 «Підприємництво, торгівля та біржова діяльність».
10. «Хмарні та GRID-технології». Робоча програма (КНТЕУ-2021). Для здобувачів вищої освіти ОС «бакалавр» спеціальності 124 «Системний аналіз», спеціальності 051 «Економіка», спеціальності 076 «Підприємництво, торгівля та біржова діяльність».
11. «Хмарні та GRID-технології». Робоча програма (ДТЕУ-2022). Для здобувачів вищої освіти ОС «бакалавр» спеціальності 124 «Системний аналіз».
12. «Хмарні та GRID-технології». Робоча програма (ДТЕУ-2022). Для здобувачів вищої освіти ОС «магістр» спеціальності 121 «Інженерія програмного забезпечення».
13. «Хмарні та GRID-технології». Збірник тестових завдань (ДТЕУ-2022). Для здобувачів вищої освіти ОС «магістр» спеціальності 121 «Інженерія програмного забезпечення» та здобувачів вищої освіти ОС «бакалавр» спеціальності 124 «Системний аналіз».
14. «Об'єктно-орієнтоване програмування». Збірник тестових завдань (ДТЕУ-2022). Для здобувачів вищої освіти ОС «бакалавр» спеціальності 121 «Інженерія програмного забезпечення».
15. Об'єктно-орієнтоване програмування». Збірник тестових завдань (ДТЕУ-2022). Для здобувачів вищої освіти ОС «бакалавр» спеціальності 125 «Кібербезпека».
16. Програма кваліфікаційного екзамену для студентів освітнього ступеня «бакалавр» (ДТЕУ-2022). Для здобувачів вищої освіти ОС «бакалавр» спеціальності 121 «Інженерія програмного забезпечення».
17. «Адміністрування та захист сховищ даних». Робоча програма (ДТЕУ-2023). Для здобувачів вищої освіти ОС «магістр» спеціальностей 121 «Інженерія програмного забезпечення» та 125 «Кібербезпека та захист інформації».

Довідку видано для подання до спеціалізованої вченої ради

**Проректор з науково-педагогічної роботи  
та міжнародних зв'язків**

**Анжеліка ГЕРАСИМЕНКО**

Божко Тетяна Василівна  
(044) 531 47 33



Довідка про участь у виконанні науково-дослідної роботи



**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**  
**ДЕРЖАВНИЙ ТОРГОВЕЛЬНО-ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ**

вул. Кіото, 19, м. Київ, 02156, тел. +380 (44) 531 47 41, e-mail: knute@knute.edu.ua, код ЄДРПОУ 44470624

29.08.2023 № 1453/24

На № \_\_\_\_\_

**ДОВІДКА**

Видана Хорольській Карині Вікторівні, здобувачу кафедри інженерії програмного забезпечення та кібербезпеки Державного торговельно-економічного університету, про те, що вона дійсно з IV кв. 2022 р. бере участь у виконанні науково-дослідної роботи «Моделювання інтелектуальних систем управління діяльністю підприємств» (термін виконання теми: I кв. 2022 р. – IV кв. 2024 р.).

Державний торговельно-економічний університет є правонаступником Київського національного торговельно-економічного університету.

Номер державної реєстрації НДР 0122U001549.

Особистий внесок Хорольської Карині Вікторівни:

- оцінено потенціал використання методів штучного інтелекту в задачах розпізнавання технічної документації підприємства;
- розроблено новий підхід до визначення інформаційної технології розпізнавання графічної інформації у технічній документації підприємства на основі використання нейронної мережі.

**Проректор**  
**з науково-педагогічної роботи**  
**та міжнародних зв'язків**



**Анжеліка ГЕРАСИМЕНКО**

Акт впровадження результатів дисертаційного дослідження на  
ТОВ «ФЬЮЧЕ КОНСАЛТИНГ»



Акт впровадження результатів дисертаційного дослідження на  
ТОВ ІНТЕРНЕТ ІНВЕСТИЦІЙНА ГРУПА



## СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА

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

23. Khorolska K., Artificial intelligence face recognition for authentication./ Kryvoruchko, O., Bebeshko, B., Khorolska, K., Desiatko, A., Kotenko, N. (2020). Technical Sciences and Technologies, 2 (20), 139-148.

*(Особистий внесок: метод отримання даних для навчання систем безпеки з соціальних мереж і взаємодії з людиною для майбутнього використання в процесі аутентифікації в різних інформаційних системах).*

24. Khorolska, K. (2022). Потенціал застосування різних методів штучного інтелекту у задачі розпізнавання креслень та трансформації 2D→3D. *Електронне фахове наукове видання "Кібербезпека: освіта, наука, техніка";*, 1(17), 21-30. <https://doi.org/10.28925/2663-4023.2022.17.2130>

25. Khorolska, K. (2022). Аналіз основних методів розпізнавання креслень та можливостей трансформації з 2D У 3D. *Електронне фахове наукове видання "Кібербезпека: освіта, наука, техніка"* 4(16), 185-193. <https://doi.org/10.28925/2663-4023.2022.16.185193>

26. Хорольська К. Аналіз основних підходів до вирішення задачі конвертації двовимірних зображень в тривимірну модель Вісник КрНУ імені Михайла Остроградського. Випуск 3/2022 (134) DOI <https://doi.org/10.32782/1995-0519.2022.3.7>

27. Bebeshko, B., Khorolska, K., Kotenko, N., Desiatko, A., Sauanova, K., Sagyndykova, S., Tyshchenko, D. 3D modelling by means of artificial intelligence (2021) Journal of Theoretical and Applied Information Technology, 99 (6), pp. 1296-1308.

*(Особистий внесок: аналіз особливостей розпізнавання 2D зображень і створення 3D моделей з використанням згортових нейронних мереж та визначенно можливі обмеження використання CNN для генерації 3D моделей з урахуванням вихідної роздільної здатності та швидкості генерації).*

28. Khorolska K., Lazorenko V., Bebeshko B., Desiatko A., Kharchenko O., Yaremych V. (2022) Usage of Clustering in Decision Support System. In: Raj J.S., Palanisamy R., Perikos I., Shi Y. (eds) Intelligent Sustainable Systems. Lecture Notes in Networks and Systems, vol 213. Springer, Singapore. [https://doi.org/10.1007/978-981-16-2422-3\\_49](https://doi.org/10.1007/978-981-16-2422-3_49)

*(Особистий внесок: побудова алгоритму на основі методу К-середніх із залежними рухами, що дозволяє конструктивно знаходити оптимальні стратегії в проектах)*

29. Lakhno, V., Akhmetov, B., Smirnov, O., Chubaievskyi, V., Khorolska, K., Bebeshko, B. (2023). Selection of a Rational Composition of Information Protection Means Using a Genetic Algorithm. In: Rajakumar, G., Du, KL., Vuppalapati, C., Beligiannis, G.N. (eds) Intelligent Communication Technologies and Virtual Mobile Networks. Lecture Notes on Data Engineering and Communications Technologies, vol 131. Springer, Singapore. [https://doi.org/10.1007/978-981-19-1844-5\\_2](https://doi.org/10.1007/978-981-19-1844-5_2)

*(Особистий внесок: побудова модуля системи підтримки прийняття рішень для вирішення задачі вибору та оптимізації кількості систем захисту інформації).*

30. Khorolska, K., Bebeshko, B., Desiatko, A., & Lazorenko, V. (2021). 3D models classification with use of convolution neural network. Paper presented at the CEUR Workshop Proceedings, 3179 25-34. [http://ceur-ws.org/Vol-3179/Paper\\_3.pdf](http://ceur-ws.org/Vol-3179/Paper_3.pdf)

*(Особистий внесок: описано мультипрезентаційну структуру класифікації 3D-моделі).*

31. Khorolska K. , Skladannyi P., Sokolov V., Korshun N., Bebeshko B., Lakhno V., Zhumadilova M (2022) Application of a convolutional neural network with a module of elementary graphic primitive classifiers in the problems of recognition of drawing documentation and transformation of 2D to 3D models. Journal of Theoretical and Applied Information Technology 31st December 2022. Vol.100. No 24 <http://www.jatit.org/volumes/Vol100No24/18Vol100No24.pdf>

*(Особистий внесок: побудовано архітектуру згорткової нейронної мережі з модулем елементарних класифікаторів графічних примітивів та розроблена модель класифікатора графічного зображення на основі охоплених класів і елементарних примітивних класифікаторів.).*

*Наукові праці, які засвідчують апробацію матеріалів дисертації:*

32. Khorolska K. Use of AI in data protection/ Kryvoruchko O., Vebeshko V., Khorolska K. // Безпека ресурсів інформаційних систем : збірник тез І Міжнародної науково-практичної конференції (м. Чернігів 16-17 квітня 2020 р.). – Чернігів : НУЧП, 2020. – с.15-18

*(Особистий внесок: описано використання технологій штучного інтелекту в основних сферах захисту даних).*

33. Бебешко Б.Т., Лазоренко В.В., Хорольська К.В. Безпека інтелектуальної системи управління цифровими активами за допомогою методу k-means при дослідженні видобутку даних // Кібергігієна. Кібербезпека. Безпека держави: матеріали наукових семінарів (Київ, 27 листопада 2020 р.)/відп. ред. АМ Десятко.–Київ: Київ. нац. торг.-екон. ун-т, 2020.–с.34-36

*(Особистий внесок: важливість використання методів нейронних мереж для виявлення закономірностей та надання їм оцінки).*

34. Лазоренко В.В., Бебешко Б.Т., Хорольська К.В. Аналіз методів прогнозування кібератак // Комплексне забезпечення якості технологічних процесів та систем (КЗЯТПС – 2021) : матеріали тез доповідей XI Міжнародної науково-практичної конференції (м. Чернігів, 26–27 травня 2021 р.) : у 2 т. / Національний університет «Чернігівська політехніка» [та ін.] ; відп. за вип.: Єрошенко Андрій Михайлович [та ін.]. – Чернігів : НУ «Чернігівська політехніка», 2021. – Т. 2. – 236 с. ISBN 978-617-7932-16-0

*(Особистий внесок: описано сьюсіб прогнозування кібератак з неповними даними за допомогою передбачення сигналу для оцінки відповідності нетрадиційних сигналів для прогнозування).*



35. Khorolska K. 3D Model reconstruction using convolutional neural networks for 2D image processing. Proceedings of the VI International Scientific and Practical Conference. Osaka, Japan. 2023. Pp.457-459 DOI: 10.46299/ISG.2023.1.6

*Наукові праці, які додатково відображають наукові результати дисертації*

36. Khorolska Karyna VR-technology as a modern architecture tool. Management of Development of Complex Systems / Tsiutsiura, Svitlana, Bebeshko, Bohdan, & Khorolska, Karyna, (2020). Management of Development of Complex Systems, 42, 69 – 74, dx.doi.org\10.32347/2412-9933.2020.42.69-74.

*(Особистий внесок: проаналізовано вплив VR на кінцевий дизайн простороої візуалізації).*

37. Хорольська К. В. UX-дизайн інформаційної системи підприємства торгівлі. / Котенко Н.О., Жирова Т.О., Десятко А.М., Хорольська К.В., Бебешко Б.Т., Тогжанова К.О. // Вісник Кременчуцького національного університету імені Михайла Остроградського. 2020. Вип. № 3 (122). С. 107–112. DOI: 10.30929/1995-0519.2020.3.67-74

*(Особистий внесок: проаналізовано інструменти для проектування клікабельного прототипу з точки зору дизайну інтерфейсу).*

38. Tetiana Zhyrova, Nataliia Kotenko, Volodymyr Tokar, Karyna Khorolska, Bohdan Bebeshko, (2021) Testing the Accessibility of Web-applications The International Scientific Journal «Computer Systems and Information Technologies» 2021, #3 DOI: <https://doi.org/10.31891/CSIT-2021-5-12>

*(Особистий внесок: здійснено ґрунтовний аналіз стандартів з питань доступності веб-додатків).*

39. Lakhno V., Akhmetov B., Ydyryshbayeva M., Bebeshko B., Desiatko A., Khorolska K. (2021) Models for Forming Knowledge Databases for Decision Support Systems for Recognizing Cyberattacks. In: Vasant P., Zelinka I., Weber

GW. (eds) Intelligent Computing and Optimization. ICO 2020. Advances in Intelligent Systems and Computing, vol 1324. Springer, Cham. [https://doi.org/10.1007/978-3-030-68154-8\\_42](https://doi.org/10.1007/978-3-030-68154-8_42)

*(Особистий внесок: описано ймовірнісні моделі для виявлення мережесих вторгнень на основі використання динамічних байесівських мереж).*

40. Lakhno, V., Mazaraki, A., Kasatkin, D., Kryvoruchko, O., Khorolska, K., Chubaievskiy, V. (2023). Models and Algorithms for Optimization of the Backup Equipment for the Intelligent Automated Control System Smart City. In: Ranganathan, G., Fernando, X., Rocha, Á. (eds) Inventive Communication and Computational Technologies. Lecture Notes in Networks and Systems, vol 383. Springer, Singapore. [https://doi.org/10.1007/978-981-19-4960-9\\_57](https://doi.org/10.1007/978-981-19-4960-9_57)

*(Особистий внесок: проаналізовано алгоритми оптимізації резервного обладнання інтелектуальної автоматизованої системи управління).*

41. Zhyrova, T., Kotenko, N., Bebeshko, B., Khorolska, K., Shevchenko, S. (2022) Benchmarking between the DQL Index and the Web Application Accessibility Index using Automatic Test Tools CEUR Workshop Proceedings, 2022, 3288, pp. 110–116

*(Особистий внесок: описано інструменти автоматизованого тестування сайтів).*

42. Khorolska K. Cyberattacks prediction with incomplete data/ Bebeshko B., Khorolska K. // Безпека соціально-економічних процесів в кіберпросторі: зб. матеріалів Всеукр. наук.-практ. конф. (Київ, 27 бер. 2019 р.). – Київ : Київ. нац. торг.-екон. ун-т, 2019. – с.123-125

*(Особистий внесок: описано інноваційні методи, такі як прогнозна імпутація сигналу (PSI), щоб замінити відсутні значення хоста, підвищуючи точність прогнозування).*

43. Khorolska K. Usage of neural networks in image recognition / O. Kryvoruchko, K. Khorolska, V. Chubaievskiy. // Зовнішня торгівля: економіка, фінанси, право. – 2019. – №3 (104). – С. 83–85.

*(Особистий внесок: класифікація частин креслення)*

44. Bebeshko, B., Khorolska, K., Kotenko, N., Kharchenko, O., & Zhyrova, T. (2021). Use of neural networks for predicting cyberattacks. Paper presented at the CEUR Workshop Proceedings, , 2923 213-223. <http://ceur-ws.org/Vol-2923/paper23.pdf>

*(Особистий внесок: описано інструмент прогнозування кібератак).*