

В. М. ПАХОМОВА

кандидат технічних наук, доцент,
доцент кафедри електронних обчислювальних машин
Дніпровський інститут інфраструктури і транспорту
Українського державного університету науки і технологій
ORCID: 0000-0002-0022-099X

К. І. ЗІНКЕВИЧ

аспірант кафедри електронних обчислювальних машин
Дніпровський інститут інфраструктури і транспорту
Українського державного університету науки і технологій
ORCID: 0009-0009-4098-9179

ОГЛЯД ТЕХНОЛОГІЙ КОМП'ЮТЕРНОГО ЗОРУ НА ОСНОВІ ГЛИБОКИХ НЕЙРОННИХ МЕРЕЖ ЩОДО ВИКОРИСТАННЯ НА ЗАЛІЗНИЧНОМУ ТРАНСПОРТІ

У даній роботі виконано огляд технологій комп'ютерного зору на основі глибоких нейронних мереж щодо використання на залізничному транспорті.

Під час воєнної агресії потребуються модернізація залізничної інфраструктури України, що пов'язана з використанням систем комп'ютерного зору на основі глибоких нейронних мереж, для забезпечення точності та безпеки. Для вирішення завдань комп'ютерного зору на залізничному транспорті (класифікація та ідентифікація об'єктів, а також стеження за ними) рекомендовано використання глибоких згорткових нейронних мереж (Deep Convolutional Neural Networks, DCNN). Проведений огляд технологій комп'ютерного зору з використанням глибоких згорткових нейронних мереж надав можливість визначити їх проблеми, відобразити загальну архітектуру з призначенням шарів (згортковий шар; шар об'єднання; активаційний шар; повноз'язаний шар; шар нормалізації; шар відсіву; цільний шар), а також ознайомитися з результатами відповідного аналізу українських та закордонних науковців відповідно до їх класифікації: на основі просторового використання; на основі глибини; з безліччю підключень; на основі використання карти ознак, стиснення і збудження; на основі використання бустінга каналів; з використанням механізму уваги. На сучасному етапі в задачах комп'ютерного зору для реалізації DCNN найчастіше використовується програмне середовище TensorFlow. Результати проведеного огляду технологій комп'ютерного зору на основі використання глибоких згорткових нейронних мереж можуть бути використані як методологічна основа для модернізації залізничної інфраструктури України в умовах воєнного стану.

Ключові слова: комп'ютерний зір, глибока згорткова нейронна мережа, проблеми, класифікація, архітектура, фільтр (ядро), функція активації, глибоке навчання, перенесення навчання, функція втрат.

V. M. PAKHOMOVA

Candidate of Technical Sciences, Associate Professor,
Associate Professor at the Department of Electronic Computers
Dnipro Institute of Infrastructure and Transport
of Ukrainian State University of Science and Technology
ORCID: 0000-0002-0022-099X

K. I. ZINKEVYCH

Postgraduate Student at the Department of Electronic Computers
Dnipro Institute of Infrastructure and Transport
of Ukrainian State University of Science and Technology
ORCID: 0009-0009-4098-9179

REVIEW OF COMPUTER VISION TECHNOLOGIES BASED ON DEEP NEURAL NETWORKS FOR USE IN RAIL TRANSPORT

This paper reviews computer vision technologies based on deep neural networks for use in railway transport.

During military aggression, modernization of the railway infrastructure of Ukraine is required, which is associated with the use of computer vision systems based on deep neural networks, to ensure accuracy and safety. To solve computer

vision tasks in railway transport (classification and identification of objects, as well as tracking them), the use of DCNN (Deep Convolutional Neural Networks) is recommended. The conducted review of computer vision technologies using deep convolutional neural networks made it possible to identify their problems, display the general architecture with the purpose of the layers (convolutional layer; pooling layer; activation layer; fully connected layer; normalization layer; dropout layer; dense layer), as well as to familiarize yourself with the results of the corresponding analysis of Ukrainian and foreign scientists according to their classification: based on spatial use; based on depth; with many connections; based on the use of a feature map, compression and excitation; based on the use of channel boosting; using the attention mechanism. At the present stage, the TensorFlow software environment is most often used in computer vision tasks to implement DCNN. The results of the review of computer vision technologies based on the use of DCNN can be used as a methodological basis for the modernization of the railway infrastructure of Ukraine under martial law.

Key words: computer vision, DCNN, problems, classification, architecture, filter (kernel), activation function, deep learning, transfer learning, loss function.

Постановка проблеми

Сучасні виклики в галузі залізничного транспорту, вимагають високого рівня автоматизації, точності та безпеки. Традиційні методи інспекції та моніторингу стану залізничної колії базуються переважно на візуальних перевірках, що вимагає значних людських ресурсів, є суб'єктивними та не завжди дозволяють вчасно виявити критичні дефекти. Окрім того, ці методи часто не здатні ідентифікувати поступові зміни, що можуть свідчити про розвиток прихованих дефектів.

Аналіз останніх досліджень і публікацій

Комп'ютерний зір (Computer Vision, CV) – галузь штучного інтелекту, пов'язана з аналізом, класифікацією і розпізнаванням зображень, відео та аудіо даних. В основі CV-систем зазвичай реалізовані алгоритми на базі машинного навчання Deep Learning, за їх допомогою вдається розпізнавати одні об'єкти від інших, відрізняти патерни та закономірності. Завдяки методам Deep Learning, CV-системи імітують властивості людського зору, розпізнають образи на зображеннях, розмічають інформацію на зображеннях, аудіо та відеоданих [2, с. 5].

Використання системи комп'ютерного зору передбачає [1, с. 48]: класифікацію об'єктів (система аналізує візуальний контент і класифікує об'єкт на фото/відео до певної категорії); ідентифікацію об'єкта (система аналізує візуальний контент та ідентифікує конкретний об'єкт на фото/відео); стеження за об'єктом (система обробляє відео, знаходить об'єкт або об'єкти, що відповідають критеріям пошуку, та відстежує його переміщення).

Для вирішення задач комп'ютерного зору можливе використання методів: контурний аналіз; метод Віюлі-Джонса; статистичні методи; приховані Марковські моделі; перетворення Карунена-Лоева; нечітка логіка; нейронні мережі та ін. На сучасному етапі доцільно використання глибоких нейронних мереж (Deep Neural Network, DNN), що доведено українськими та закордонними вченими в наукових працях.

У [2, с. 6] подано класифікацію DNN, яка передбачає: навчання з вчителем (багатосаровий перцептрон – Multilayer Perceptron, MLP; згорткові нейронні мережі – Convolutional Neural Networks, CNN); навчання без вчителя (глибокі мережі довіри – Deep Belief Network, DBN; рекурентні нейронні мережі – Recurrent Neural Networks, RNN); гібридне навчання (генеративно-конкурентні мережі – Generative Adversarial Networks, GAN; рекурсивні нейронні мережі – Recursive Neural Networks, RvNN). Останнім часом, для вирішення задач комп'ютерного зору в різних сферах, широко використовуються такі моделі глибокого навчання, які прийнято називати згортковими нейронними мережами.

З одного боку, CNN схожі з звичайними нейронними мережами, з іншого боку, вони спроектовані з припущенням, що на вхід подається зображення. На відміну від звичайних нейронних мереж, шари CNN є тривимірними. Нейрони одного шару пов'язані тільки з певною областю наступного шару. Перший вхідний шар мережі має розмірність $w \times h \times d$, де w – ширина зображення, h – висота, d – кількість каналів кольору [4, с. 53].

Формулювання мети дослідження

Мета дослідження є аналіз можливостей технологій комп'ютерного зору на основі глибоких згорткових нейронних мереж щодо використання в сфері автоматизації моніторингу та діагностики елементів залізничної інфраструктури. Відповідно до мети поставлені наступні завдання: визначити проблеми глибоких нейронних мереж; вивчити архітектуру та класифікацію згорткових нейронних мереж; ознайомитися з використанням систем комп'ютерного зору на залізничному транспорті.

Викладення основного матеріалу дослідження

Проблеми глибоких нейронних мереж. Зінченко О. В., Звенігородський О. С., Кисіль Т. М. довели, що із збільшенням глибини нейронної мережі точність обробки на початку збільшується, а в подальшому значно погіршується. Проте, зниження точності навчання ускладнює процес оптимізації. В процесі підвищення поточних характеристик CNN можна використовувати в процесі їх моделювання наступні методи: мінімізацію при зменшенні розмірності, нормалізацію батчів, збільшення та розширення вибірки, просунуті режими оптимізації, застосування різних типів функцій активації [2, с. 11].

Основною складністю роботи CNN є тривале навчання, яке може складати від кількох днів до місяців [3, с. 64]. Крім того, для успішного навчання потрібна велика кількість розмічених даних. У зв'язку з цим крім побудови нейронної мережі з нуля (from scratch) активно застосовується підхід перенесення навчання (transfer learning). Для DNN використовується підходи до перенесення навчання: метод, що ґрунтується на моделі (model-based); метод, що ґрунтується на ознаках (feature-based) [3, с. 64].

Мамута М. С., Кравченко І. В., Мамута О. Д., Тужанський С. Є. акцентують увагу, що основна проблема, яка характерна для CNN (як для всіх DNN) – це проблема затухання градієнта. Причиною цього є те, що на початковому етапі вагові коефіцієнти задаються за допомогою генератора випадкових чисел, а отже вони ще є «поганими», і лише в процесі навчання ці коефіцієнти набувають сенсу [3, с. 65]. Ще однією складністю при роботі з CNN є необхідність великої кількості даних. Звичайно ж є готові набори, ще й розмічені, інколи навіть з результатами роботи стандартних алгоритмів. У такому разі зручно перевірити та порівняти роботу створеної моделі з існуючими. Але складнощі виникають при вузькоспеціалізованих задачах, оскільки набори даних зазвичай відсутні, принаймні у відкритому доступі [3, с. 66]. Слід зауважити, що існують відкриті бази зображень та відео, на яких розробники алгоритмів порівнюють свої напрацювання; такими базами є CIFAR-10, CIFAR-100, ILCVRC (ImageNet Large Scale Visual Recognition Challenge), PASCAL VOC Challenge, COCO (Common Objects in Context) [4, с. 54].

Архітектура згорткової нейронної мережі. У загальному випадку архітектура DCNN складається з таких шарів [8, с. 1]: згортковий шар; шар об'єднання; активаційний шар; повнозв'язаний шар; шар нормалізації; шар відсіву; щільний шар.

Згортковий шар (convolution layer) відповідає за вилучення ознак із вхідного зображення; він виконує операцію згортки над вхідним зображенням, де до зображення застосовується фільтр або ядро для ідентифікації та вилучення певних ознак.

Шар об'єднання (pooling layer) відповідає за зменшення просторових розмірів карт ознак, створених згортковим шаром; він виконує операцію зниження розміру карт ознак, щоб зменшити розмір карт ознак та зменшити обчислювальну складність.

Активаційний шар (activation layer) застосовує нелінійну активаційну функцію, таку як функція ReLU, до виходу шару об'єднання; ця функція допомагає ввести нелінійність у модель, дозволяючи їй вивчати складніші представлення вхідних даних.

Повнозв'язаний шар (fully connected layer) – це традиційний шар нейронної мережі, який з'єднує всі нейрони попереднього шару з усіма нейронами наступного шару. Цей шар відповідає за об'єднання ознак, вивчених згортковим та об'єднувальним шарами, для створення прогнозу.

Шар нормалізації (normalization layer) виконує операції нормалізації, такі як пакетна нормалізація або нормалізація шарів, щоб забезпечити належну обумовленість активацій кожного шару та запобігти перенаванчання.

Шар відсіву (dropout layer) використовується для запобігання перенаванчання шляхом випадкового відсіву нейронів під час навчання. Це допомагає гарантувати, що модель не запам'ятовує навчальні дані, а натомість узагальнює їх на нові, невидимі дані.

Щільний шар (dense layer): після того, як згортковий та об'єднувальний шари витягли ознаки з вхідного зображення, щільний шар можна використовувати для об'єднання цих ознак та створення остаточного прогнозу. У CNN щільний шар зазвичай є останнім шаром і використовується для створення вихідних прогнозів. Активності з попередніх шарів зрівнюються та передаються як вхідні дані до щільного шару, який виконує зважену суму вхідних даних та застосовує функцію активації для створення остаточного вихідного сигналу.

У якості прикладу на рис. 1 [8] показано архітектуру CNN.

Томка Ю. Я., Талах М. В., Дворжах В. В., Ушенко О. Г. зазначають, що кожен шар отримує об'ємну 3D-інформацію на вході та перетворює її з відповідним зберіганням 3D-об'єму за допомогою диференційованої функції [10, с. 57]. Науковці розглянули узагальнений алгоритм реалізації CNN у програмному середовищі TensorFlow для класичної задачі класифікації розрядів набору даних MNIST.

Важливо, що DCNN здатна досягати рекордних результатів на дуже складному наборі даних, використовуючи виключно контрольоване навчання; крім того, продуктивність DCNN знижується, якщо видалити один згортковий шар [7, с. 90].

Класифікація згорткових нейронних мереж: CNN на основі просторового використання (LeNet, AlexNet, ZfNet, VGG, GoogleNet); CNN на основі глибини (Highway Net – магістральна, ResNet – залишкова, Inception V3, Inception V4, Inception ResNet); розгалужені CNN (Highway Net, ResNet, DenseNet); CNN з безліччю підключень (WideResNet, Pyramidal Net, Xception, Inception Family, ResNet); CNN на основі використання карти ознак (Squeeze and Excitation), стиснення і збудження (Competitive Squeeze and Excitation); CNN на основі використання бустінга каналів (Channel Boosted CNN); CNN з використанням механізму уваги (Residual Attention Neural Network, Convolution Block Attention, Concurrent Squeeze, Excitation).

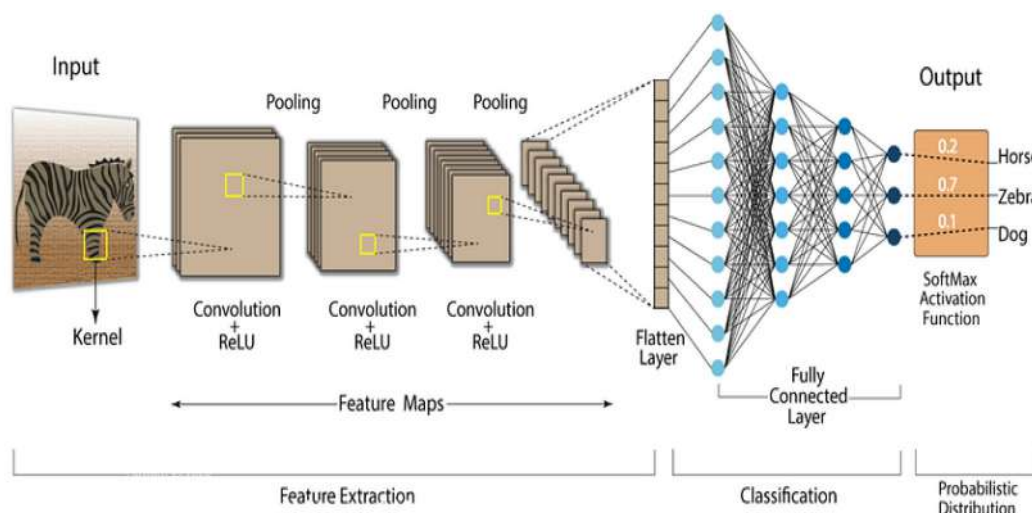


Рис. 1. Архітектура CNN [8]

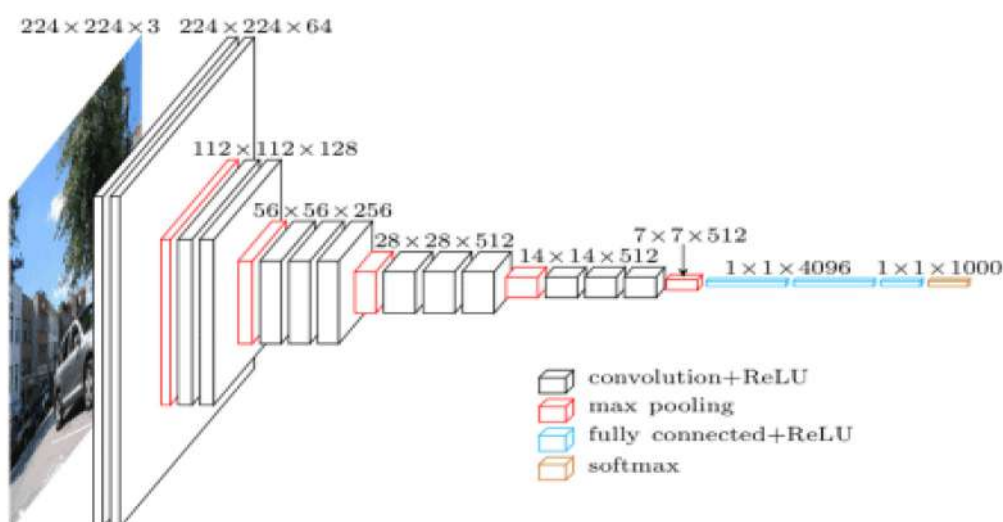


Рис. 2. Архітектура VGG [5]

Як в Україні, так і за її межами різними вченими і науковцями проведений відповідний аналіз CNN в наступних роботах: [2, с. 10–11]; [3, с. 66]; [4, с. 54]; [10, с. 55–56]. Так, наприклад, особливістю GoogleNet є використання декількох фільтрів різних розмірів на кожному згортковому шарі (технологія inception) [3, с. 66]. Таким чином, карта ознак одночасно сканується декількома фільтрами, які виділяють різну інформацію з вхідних даних. Далі здійснюється конкатенація отриманих результатів і передача в наступний шар. Для боротьби в GoogleNet із затуханням градієнтів використовуються додаткові класифікатори, тобто блок повнозв'язних шарів розташовується не лише після всіх згорткових шарів, а й в декількох місцях всередині CNN. Функція втрат (loss) рахується для всіх класифікаторів і для алгоритму зворотнього поширення помилки використовується складний loss. За рахунок цього вся CNN ефективно навчається [3, с. 66].

У якості прикладу на рис. 2 [5, с. 1] показано VGG – це архітектура класичної DCNN.

Використання систем комп'ютерного зору на залізничному транспорті. У результаті аналізу встановлено, що системи комп'ютерного зору, поєднані з DNN, здатні аналізувати відео та зображення в реальному часі, виявляючи й класифікуючи різні об'єкти: залізничні вагони; номери локомотивів; бар'єри та ін., проводити автоматичний огляд стану залізничної колії. Нейронні мережі DNN здатні виявляти навіть незначні тріщини, відхилення рівня або зношення елементів інфраструктури, які важко помітити людським оком, що в свою чергу значно знижує ризик аварій. Для реалізації такої системи необхідне поєднання тривимірного машинного зору з технологією автоматичного виявлення змін. Зміни визначаються за допомогою DCNN, яка дозволяє точно порівнювати стан колії між різними інспекційними проходами [6, с. 144].

У [9, с. 49] дослідження зосереджено на розробці високопродуктивної системи інспекції залізниць, адаптованої для високошвидкісних залізниць та залізниць з обмеженими часовими рамками інспекції завдяки інтеграції методів глибокого навчання та багатоядерних обчислювальних платформ.

Використання нейронних мереж в гальмівних системах дозволяє зменшити гальмівний шлях поїздів, підвищити безпеку та знизити витрати на обслуговування. Такі системи можуть аналізувати дані про швидкість, вагу поїзда та стан колії, щоб оптимізувати силу гальмування, контролювати правильність розміщення та закріплення вантажу, виявляючи потенційно небезпечні зміщення під час руху. У деяких країнах застосовуються подібні технології, наприклад, у Німеччині та Японії використовують дрони з комп'ютерним зором для інспекції колій, а в Китаї запроваджено систему автоматичного моніторингу поїздів із застосуванням нейромережових технологій, однак в Україні відбуваються перші кроки в цьому напрямку.

Висновки

Технології комп'ютерного зору з використанням DCNN відкривають нові можливості для модернізації залізничної інфраструктури. Особливо перспективним є використання DCNN у поєднанні з 3D аналізом, що дозволяє відстежувати важливі проблеми з високою точністю. Подальші дослідження мають бути зосереджені на адаптації CV-систем з використанням DCNN в залізничній мережі України.

Список використаної літератури

1. Гангало І.М., Лісовий Д.О., Жебка В.В. Розпізнавання об'єктів за допомогою технологій комп'ютерного зору. *Телекомунікаційні та інформаційні технології*. 2022. № 4(77). С. 46–52. DOI: 10.31673/2412-4338.2022.044652
2. Зінченко О. В., Звенігородський О. С., Кисіль Т. М. Згорткові нейронні мережі для вирішення задач комп'ютерного зору. *Телекомунікаційні та інформаційні технології*. 2022. № 2(75). С. 4–12. DOI: 10.31673/2412-4338.2022.020411.
3. Мамута М. С., Кравченко І. В., Мамута О. Д., Тужанський С. Є. Оцінка класифікації зображень для перенесення навчання у згорткових нейронних мережах. *Оптико-електронні інформаційно-енергетичні технології*. 2023. № 1(45). С. 64–69. DOI: 10.31649/1681-7893-2023-45-1-64-70
4. Тимчишин Р. М., Волков О. Є., Господарчук О. Ю., Богачук Ю. П. Сучасні підходи до розв'язання задач комп'ютерного зору. *Control systems and computers*, 2018. № 6. С. 46–73. DOI: <https://doi.org/10.15407/usim.2018.06.046>
5. Boesch G. Very Deep Convolutional Networks (VGG) Essential Guide. 2021. [Electronic resource]. URL: <https://viso.ai/deep-learning/vgg-very-deep-convolutional-networks/> (дата звернення: 20.08.2025)
6. Harrington R. M., Lima A. de O., Fox-Ivey R., Nguyen T., Laurent J., Dersch M. S., Edwards J. R.. Use of deep convolutional neural networks and change detection technology for railway track inspections. *Journal of Rail and Rapid Transit*. 2023. Vol. 237. Iss. 2. Pp. 137–145. URL: <https://doi.org/10.1177/09544097221093486>
7. Krizhevsky A., Sutskever I., Hinton G.E. ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*. 2017. Vol. 60. No. 6. Pp.84-90. URL: <https://doi.org/10.1145/3065386>
8. Nafizul Haque Kh. What is Convolutional Neural Network – CNN (Deep Learning). 2023. [Electronic resource]. URL:<https://nafizshahriar.medium.com/what-is-convolutional-neural-network-cnn-deep-learning-b3921bdd82d5> (дата звернення: 20.08.2025).
9. Tang Y., Qian Yu. High-speed railway track components inspection framework based on YOLOv8 with high-performance model deployment. *High-speed Railway*. 2024. № 2(1). Pp. 42–50. URL: <https://doi.org/10.1016/j.hspr.2024.02.001>
10. Tomka Yu., Talakh M., Dvorzhak V., Ushenko O. Implementation of a Convolutional Neural Network Using TensorFlow Machine Learning Platform. *Optoelectronic Information-Power Technologies*. 2023. Vol. 44. № 2. Pp. 55–65.

References

1. Hanhalo I. M., Lisovyy D. O., Zhebka V. V. (2022). Rozpiznavannya ob'ektiv za dopomohoyu tekhnolohiy komp'yuternoho zoru [Object recognition using computer vision technologies]. *Telecommunication and information technologies*, 4(77), 46–52. DOI: 10.31673/2412-4338.2022.044652 [in Ukrainian].
2. Zinchenko O. V., Zvenihorodskyy O. S., Kysil T. M. (2022). Z-hortkovi neyronni merezhi dlva vyrishennya zadach komp'yuternoho zoru [Convolutional neural networks for solving computer vision problems]. *Telecommunication and information technologies*, 2(75), 4–12. DOI:10.31673/2412-4338.2022.020411 [in Ukrainian].
3. Mamuta M. S., Kravchenko I. V., Mamuta O. D., Tuzhans'kyi S. Ye. (2023). Otsinka klasyfikatsiyi zobrazhen dlya perenesennya navchannya u z-hortkovykh neyronnykh merezhakh [Image classification assessment for transfer learning in convolutional neural networks]. *Optoelectronic Information-Power Technologies*, 1(45), 64–69. DOI: 10.31649/1681-7893-2023-45-1-64-70 [in Ukrainian].

4. Tymchyshyn R. M., Volkov O. Ye., Hospodarchuk O. Yu., Bohachuk Yu. P. (2018). Suchasni pidkhody do rozv'yazannya zadach komp'yuternoho zoru [Modern approaches to solving computer vision problems]. *Control systems and computers*, 6, 46–73. DOI: <https://doi.org/10.15407/usim.2018.06.046> [in Ukrainian].
5. Boesch G. (2021). Very Deep Convolutional Networks (VGG) Essential Guide [Electronic resource]. URL: <https://viso.ai/deep-learning/vgg-very-deep-convolutional-networks/> [in English].
6. Harrington R. M., Lima A. de O., Fox-Ivey R., Nguyen T., Laurent J., Dersch M. S., Edwards J. R. (2023). Use of deep convolutional neural networks and change detection technology for railway track inspections. *Journal of Rail and Rapid Transit*, 237, 2, 137–145. URL: <https://doi.org/10.1177/09544097221093486> [in English].
7. Krizhevsky A., Sutskever I., Hinton G. E. (2017). ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, 60, 6, 84–90. URL: <https://doi.org/10.1145/3065386> [in English].
8. Nafizul Haque Kh. What is Convolutional Neural Network – CNN (Deep Learning). (2023). [Electronic resource]. URL: <https://nafizshahriar.medium.com/what-is-convolutional-neural-network-cnn-deep-learning-b3921bdd82d5> [in English].
9. Tang Y., Qian Yu. (2024). High-speed railway track components inspection framework based on YOLOv8 with high-performance model deployment. *High-speed Railway*, 2(1), 42–50. URL: <https://doi.org/10.1016/j.hspr.2024.02.001> [in English].
10. Tomka Yu., Talakh M., Dvorzhak V., Ushenko O. (2023). Implementation of a Convolutional Neural Network Using TensorFlow Machine Learning Platform. *Optoelectronic Information-Power Technologies*, 44, 2, 55–65. DOI: <https://doi.org/10.31649/1681-7893-2022-44-2-55-65> [in English].

Дата першого надходження рукопису до видання: 15.09.2025

Дата прийнятого до друку рукопису після рецензування: 13.10.2025

Дата публікації: 28.11.2025