УДК 004.932:681.3

# ЗАСТОСУВАННЯ КОМП'ЮТЕРНОГО ЗОРУ ДЛЯ РОЗПІЗНАВАННЯ ВІЗУАЛЬНИХ ОБ'ЄКТІВ У ЦИФРОВОМУ СЕРЕДОВИЩІ

***[1]Олександр СУПРУН, [2]Олександра КАЛІНІНА***
*[1] канд. техн. наук, доцент, кафедра МСТ, ХНУРЕ*
*[2]студентка, кафедра МСТ, ХНУРЕ*

***Abstract.*** *The paper examines the application of computer vision technologies for visual object recognition in digital environments. Modern methods of image processing and feature extraction are analyzed. Particular attention is paid to machine learning and deep neural networks used for classification tasks. The practical effectiveness of computer vision systems in multimedia, security, industrial automation and intelligent information systems is evaluated.*

*Keywords: computer vision, image recognition, neural networks, object detection, digital environment.*

Комп'ютерний зір є одним із найбільш перспективних напрямів розвитку сучасних інформаційних технологій [1]. Його основне завдання полягає в автоматизованому отриманні, аналізі та інтерпретації візуальної інформації, представленої у вигляді цифрових зображень або відеопотоків. Активний розвиток алгоритмів машинного навчання та нейронних мереж значно підвищив точність розпізнавання об'єктів у складних умовах цифрового середовища. Системи комп'ютерного зору широко застосовуються у відеоспостереженні, медичній діагностиці, автономному транспорті, робототехніці, промисловому контролі якості продукції та мультимедійних системах. Основною перевагою таких систем є можливість автоматичного виявлення та класифікації об'єктів у режимі реального часу [2, 3].

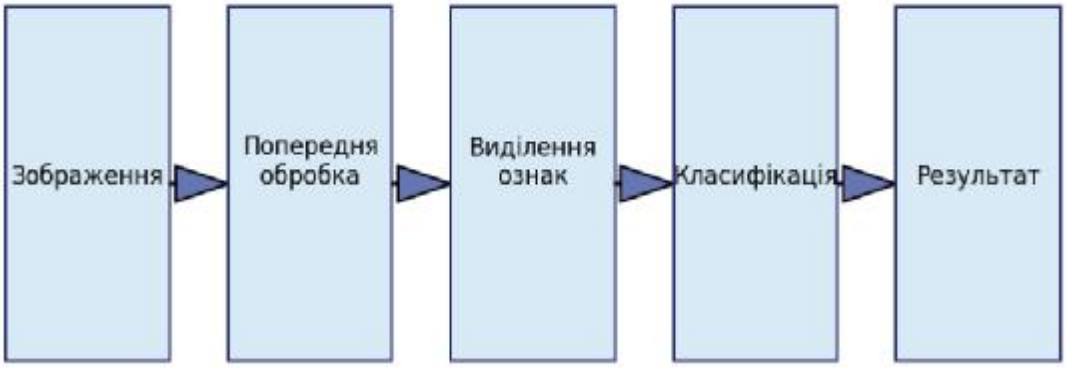


Рис. 1. Основні етапи розпізнавання візуальних об'єктів у системах комп'ютерного зору

Процес розпізнавання візуальних об'єктів складається з декількох послідовних етапів: отримання зображення, попередньої обробки, виділення інформативних ознак, класифікації та формування результату (рис. 1). Для підвищення точності використовуються згорткові нейронні мережі, які дозволяють автоматично формувати набір релевантних ознак без додаткового втручання оператора [4].

Для оцінювання ефективності розпізнавання використовується показник точності класифікації *Accuracy*:

$$Accuracy = N_{correct} / N_{total} \times 100\ \%, \quad (1)$$

де $N_{correct}$ – кількість правильно розпізнаних об’єктів; $N_{total}$ – загальна кількість об’єктів тестової вибірки.

Використання даного критерію дозволяє порівнювати різні алгоритми розпізнавання та визначати їх придатність для конкретних прикладних задач. Практичні дослідження показують, що сучасні згорткові нейронні мережі забезпечують точність понад 95 % для багатьох типів візуальних даних. Подальший розвиток комп’ютерного зору пов’язаний із застосуванням генеративного штучного інтелекту, мультимодальних моделей та технологій оброблення великих масивів даних. Це створює передумови для впровадження інтелектуальних систем нового покоління, здатних виконувати складний аналіз цифрового контенту.

Таким чином, технології комп’ютерного зору є важливим інструментом автоматизації процесів аналізу зображень та відео. Їх використання забезпечує підвищення швидкості обробки інформації, зменшення впливу людського фактора та розширення функціональних можливостей цифрових систем. Завдяки застосуванню сучасних згорткових нейронних мереж досягається висока точність класифікації об’єктів у режимі реального часу. Подальше інтегрування генеративного штучного інтелекту та мультимодальних моделей відкриває широкі перспективи для створення інтелектуальних систем наступного покоління.

**Список літератури**

1. Barua, A., Ahmed, M. U., & Begum, S. (2023). A systematic literature review on multimodal machine learning: Applications, challenges, gaps and future directions. *IEEE Access*, *11*, 14804–14831. https://doi.org/10.1109/ACCESS.2023.3243854

2. Karlov, D. V., Tupitsya, I. M., & Parkhomenko, M. V. (2022). Methodology of increasing the reliability of video information in infocommunication networks aerosegment. *Radio Electronics, Computer Science, Control*, *(3)*, 120. https://doi.org/10.15588/1607-3274-2022-3-12

3. Stahlschmidt, S. R., Ulfenborg, B., & Synnergren, J. (2022). Multimodal classification: Current landscape, taxonomy and future directions. *ACM Computing Surveys*, *54*(10s), 1–38. https://doi.org/10.1145/3543848

4. Wang, B., Feng, Y., Xiong, X.-C., Qiang, Y.-H., & Qiang, B.-H. (2023). Multi-modal transformer using two-level visual features for fake news detection. *Applied Intelligence*, *53*(9), 10429–10443. https://doi.org/10.1007/s10489-023-04548-6