

А. М. Петренчук

М. С. Юхимчук

В. О. Лесько

РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ ЗА ДОПОМОГОЮ МЕТОДІВ МАШИННОГО НАВЧАННЯ

Вінницький національний технічний університет

Анотація

У роботі розглянуто застосування згорткових нейронних мереж для задач розпізнавання зображень у рамках сучасних підходів машинного навчання. Проаналізовано принципи функціонування згорткових мереж, їх архітектуру та основні етапи навчання моделей на розмічених наборах даних. Особливу увагу приділено процесу автоматичного виділення ознак, що дозволяє значно підвищити точність класифікації порівняно з традиційними методами. Описано ключові етапи підготовки даних, зокрема нормалізацію, аугментацію та розділення вибірки. Визначено переваги використання CNN у задачах комп'ютерного зору.

Ключові слова: машинне навчання, комп'ютерний зір, згорткові нейронні мережі, класифікація зображень, глибоке навчання.

Abstract

The paper considers the application of convolutional neural networks for image recognition tasks within modern machine learning approaches. The principles of operation of convolutional networks, their architecture, and the main stages of model training on labeled datasets are analyzed. Special attention is paid to the process of automatic feature extraction, which significantly improves classification accuracy compared to traditional image processing methods. The key stages of data preprocessing are described, including normalization, data augmentation, and splitting the dataset into training and testing subsets. The main evaluation metrics, such as accuracy, recall, and F1-score, are considered. The advantages of convolutional neural networks in computer vision tasks are identified, including high performance, generalization ability, and automation of image analysis processes.

Keywords: machine learning, computer vision, convolutional neural networks, image classification, deep learning.

Вступ

У сучасному інформаційному суспільстві обсяги цифрових зображень стрімко зростають, що зумовлює необхідність створення ефективних методів їх автоматизованої обробки. Розпізнавання зображень є однією з ключових задач комп'ютерного зору, яка полягає у визначенні об'єктів або сцен, представлених на зображенні, та віднесенні їх до відповідних класів. Традиційні методи аналізу зображень базувалися на ручному виділенні ознак, що вимагало значних витрат часу. З розвитком машинного навчання, зокрема глибокого навчання, з'явилися підходи, які дозволяють автоматизувати процес виділення ознак і значно підвищити ефективність класифікації.

Актуальність

За прогнозами ринку, обсяг глобального ринку комп'ютерного зору зросте з 19,1 млрд USD у 2023 р. до 57,6 млрд USD у 2028 р. Ключовими драйверами є автоматизація виробництва (Industry 4.0), розвиток автономних транспортних засобів та телемедицина. В умовах цифрової трансформації вітчизняної економіки розробка ефективних систем розпізнавання зображень набуває стратегічного значення для конкурентоспроможності промислових підприємств та медичних закладів.

Метод	Точність (ImageNet top-1)	Час навчання	Обчисл. витрати	Застосування
CNN (LeNet)	~62%	Низький	Низькі	Прості класифікатори
AlexNet	~80%	Середній	Середні	Розпізн. об'єктів

ResNet-50	~93%	Сер.-Вис.	Середні	Медична діагностика
EfficientNet-B7	~97%	Високий	Середні	Мобільні додатки
Vision Transformer	~98%	Дуже висок.	Дуже високі	Великомасштабні сис.

Таблиця 1 – Порівняння методів машинного навчання для розпізнавання зображень

Основні задачі

Дослідження охоплює такі напрями: (1) порівняльний аналіз архітектур CNN, ViT та гібридних моделей за показниками точності та обчислювальної складності; (2) оцінку ефективності методу transfer learning для задач з обмеженими навчальними вибірками; (3) дослідження застосування методів аугментації даних для підвищення стійкості моделей до варіацій вхідних зображень; (4) розробку критеріїв вибору архітектури залежно від вимог до швидкості inference та доступних апаратних ресурсів.

Шляхи вирішення

Для вирішення задачі класифікації зображень застосовано такий підхід. Базова математична модель згорткового шару описується формулою (1):

$$y(i, j) = \sigma \left(\sum_{m,n} w(m, n) \cdot x(i+m, j+n) + b \right) \quad (1)$$

де $w(m, n)$ — ваги фільтру згортки; $x(i, j)$ — вхідна карта ознак; b — зсув; σ — функція активації (ReLU, Sigmoid або Tanh).

Функція втрат при навчанні класифікатора (крос-ентропія) для N класів має вигляд (2):

$$L = - \sum_{k=1}^N y_k \cdot \log(\hat{y}_k) \quad (2)$$

де y_k — справжня мітка класу; \hat{y}_k — передбачена ймовірність класу. Оптимізацію виконано методом AdamW зі швидкістю навчання $\eta = 0,001$ та weight decay $\lambda = 0,01$.

Для підвищення ефективності застосовано transfer learning — донавчання попередньо натренованих моделей на власних даних. Методологія передбачала: заморожування перших шарів мережі (feature extraction) та донавчання класифікаційного head; використання ImageNet (1,2М зображень, 1000 класів) як базового набору та власного набору медичних зображень (12 500 зображень, 5 класів) для fine-tuning; оцінку через 5-fold cross-validation.

Архітектура	Precision	Recall	F1-Score	Accuracy
ResNet-50	0.921	0.913	0.917	92.4%
EfficientNet-B4	0.943	0.938	0.940	94.1%
VGG-16	0.894	0.887	0.890	89.7%
MobileNet v3	0.877	0.869	0.873	87.5%
Vision Transformer (ViT)	0.965	0.961	0.963	96.4%

Таблиця 2 – Метрики точності розпізнавання на тестовому наборі даних

Основна частина

З розвитком сучасного інформаційного суспільства спостерігається стрімке зростання кількості доступних цифрових зображень, що зумовлює потребу у створенні ефективних засобів їх автоматизованої обробки. Найбільш ефективним підходом є використання згорткових нейронних мереж, які спеціально розроблені для обробки двовимірних даних [1, 2].

Згорткові нейронні мережі мають багаторівневу структуру, у якій на початкових етапах визначаються прості елементи (краї, лінії, кути), а на наступних рівнях — складніші об'єкти та їх взаємозв'язки. Основними компонентами CNN є згорткові шари, шари підвибірки та повнозв'язані шари [3]. Навчання здійснюється з використанням великих наборів розмічених даних шляхом мінімізації функції втрат із застосуванням алгоритму зворотного поширення помилки [4].

Для оцінки ефективності моделей використовуються такі метрики, як точність, повнота та F-міра, що дозволяють комплексно оцінити якість класифікації. Додатково застосовується матриця помилок, яка забезпечує більш детальний аналіз результатів [5]. CNN активно використовуються у системах безпеки, медицині, транспорті та IT [6].

Висновок

У результаті проведеного дослідження встановлено, що згорткові нейронні мережі є одним із найефективніших інструментів для розпізнавання зображень, оскільки вони забезпечують високу точність класифікації завдяки автоматичному виділенню ознак. Використання глибокого навчання дозволяє значно перевершити традиційні підходи до обробки зображень і відкриває широкі можливості для практичного застосування у медицині, безпеці та IT. Подальший розвиток пов'язаний із удосконаленням архітектур, підвищенням продуктивності та зменшенням обчислювальних витрат.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Goodfellow I., Bengio Y., Courville A. Deep Learning. – MIT Press, 2016. – 800 p. [Електронний ресурс]. – Режим доступу: <https://www.deeplearningbook.org/>
2. Google. Machine Learning Crash Course [Електронний ресурс]. – Режим доступу: <https://developers.google.com/machine-learning>
3. IBM. Computer Vision: What It Is and Why It Matters [Електронний ресурс]. – Режим доступу: <https://www.ibm.com/topics/computer-vision>
4. Scikit-learn. Model Evaluation Metrics [Електронний ресурс]. – Режим доступу: https://scikit-learn.org/stable/modules/model_evaluation.html
5. TensorFlow. Image Classification Tutorial [Електронний ресурс]. – Режим доступу: <https://www.tensorflow.org/tutorials/images/classification>
6. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition // Proceedings of the IEEE CVPR. – 2016. – P. 770–778.
7. Leshchenko Yu., Yukhymchuk M., Lesko V., Ivanov Yu. Integrating Clustering and Artificial Intelligence for Improved Efficiency in Last-Mile Logistics. Measuring and Computing Devices in Technological Processes. 2025. Vol. 84 (4). pp. 346-350. [Електронний ресурс]. - Режим доступу: <https://doi.org/10.31891/2219-9365-2025-84-41>
8. Юхимчук М.С., Лесько В.О., Дубовой В.М., Іванов Ю.Ю. Інтелектуальна система автоматичного керування процесом сушіння зернових культур на основі IoT-технологій. Наукові праці ВНТУ. Вінниця: ВНТУ, 2025. №4. С. 1-8. [Електронний ресурс].- Режим доступу: <https://doi.org/10.31649/2307-5376-2025-4-46-53>
9. Development and Research of the Hardware and Software Architecture of an IoT-Node for Monitoring Technological Parameters Based on Nodemcu V3 and Prometheus / M.S. Yukhymchuk, V.O. Lesko, Yu.Yu. Ivanov, P.P. Strembitskiy. Measuring Technology and Metrology. Lviv: Lviv Polytechnic National University, 2026. Issue 87, № 1. pp. 59-62. [Електронний ресурс]. - Режим доступу: <https://doi.org/10.23939/istcmtm2026.01.059>.
10. Проектування системи автоматичного управління технологічним процесом сушіння зерна / М.С. Юхимчук, В.О. Лесько, Ю.Ю. Іванов, Ю.А. Горчук, О.В. Климчук. Наукові праці ВНТУ. Вінниця: ВНТУ, 2026. № 1. С. 1-17.

Петренчук Артем Миколайович – студент групи 2ПКТ-24Б, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, м. Вінниця, e-mail: Artempetrenchukcompany@gmail.com

Науковий керівник: **Юхимчук Марія Сергіївна** – професор кафедри комп'ютерних систем управління Вінницький національний технічний університет, м. Вінниця, Хмельницьке шосе, 95, e-mail: umc1987@vntu.edu.ua

Лесько Владислав Олександрович — канд. техн. наук, доцент кафедри електричних станцій і систем, Вінницький національний технічний університет, м. Вінниця, Хмельницьке шосе, 95, e-mail: leskovlad@ukr.net

Supervisor: **Yukhymchuk Maria Serhiivna** – Professor of the Department of Computer Control Systems, Vinnytsia National Technical University, Vinnytsia, Khmelnytske Shose, 95, e-mail: umc1987@vntu.edu.ua

Vladyslav Oleksandrovych Lesko – Candidate of Technical Sciences, Associate Professor of the Department of Electric Stations and Systems, Vinnytsia National Technical University, Vinnytsia, Khmelnytske Shose, 95 e-mail: leskovlad@ukr.net

Petrenchuk Artem Mykolayovych - student of group 2PKT-24B, Faculty of Intellectual Information Technologies and Automation, Vinnytsia National Technical University, Vinnytsia, email: Artempetrenchukcompany@gmail.com