

УДК: 004.8

*Зелінський О. О, здобувач вищої освіти,
Січко Т. В., канд. техн. наук, доцент,
доцент кафедри інформаційних технологій,
Донецький національний університет імені Василя Стуса*

ЗАСТОСУВАННЯ МЕТОДІВ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ІДЕНТИФІКАЦІЇ ЗОБРАЖЕНЬ

Анотація. У роботі розглянуто сучасні методи ідентифікації зображень на основі штучного інтелекту. Проаналізовано використання згорткових нейронних мереж і трансформерів, а також визначено перспективи розвитку систем розпізнавання об'єктів у різних сферах діяльності.

Ключові слова: штучний інтелект, комп'ютерний зір, нейронна мережа, ідентифікація, трансформер.

Вступ. Ідентифікація зображень є одним із ключових напрямів розвитку штучного інтелекту та комп'ютерного зору. Зростання обчислювальних потужностей, розвиток нейронних мереж і доступність великих обсягів даних дали змогу створювати системи, здатні автоматично розпізнавати об'єкти, сцени й події з високою точністю.

Сучасні методи машинного навчання дають змогу моделювати складні залежності у візуальних даних і виявляти закономірності, недоступні для класичних алгоритмів обробки зображень. Це зумовлює актуальність дослідження застосування методів штучного інтелекту в завданнях ідентифікації зображень, а також необхідність аналізу сучасних архітектур і практичних підходів у цій сфері.

1. Теоретичні засади ідентифікації зображень

Ідентифікація зображень полягає у виділенні значущих ознак з візуального контенту та подальшому їх порівнянні з еталонними шаблонами або категоріями. Традиційно ця задача вирішувалася з використанням алгоритмів комп'ютерного зору – виділення контурів, ключових точок, гістограм напрямків градієнтів (HOG), або методів на основі дескрипторів ознак (SIFT, SURF). Проте такі підходи мали обмеження у стійкості до змін освітлення, масштабу й ракурсу [1].

Із розвитком штучного інтелекту відбувся перехід до використання глибоких нейронних мереж (Deep Neural Networks, DNN), які здатні автоматично вилучати релевантні ознаки із сирих зображень без ручного проектування. Найбільш ефективними виявилися згорткові нейронні мережі (Convolutional Neural Networks, CNN), які обробляють дані у вигляді двовимірних масивів і вчать виявляти просторові структури [2].

Завдяки цьому вдалося значно підвищити точність класифікації, локалізації та сегментації об'єктів. Класичні архітектури, такі як AlexNet, VGG, ResNet і Inception, стали базовими у побудові систем розпізнавання зображень [2].

2. Сучасні архітектури та підходи до побудови моделей

Поява глибоких мереж із залишковими зв'язками (ResNet) дала змогу навчати моделі, що складаються з десятків і сотень шарів, без втрати точності. Ви-

користання блоку skip connection дало змогу уникнути проблеми зникання градієнта та покращити збіжність навчання [3].

Мережі типу Insertion поєднують у межах одного шару фільтри різного розміру, що підвищує ефективність вилучення ознак різних масштабів. Подальші вдосконалення, зокрема EfficientNet, забезпечили оптимальне співвідношення між точністю, кількістю параметрів і швидкістю моделі [3].

У наступних поколіннях моделей з'явилися візуальні трансформери (Vision Transformers, ViT), які базуються на механізмі самоуваги (self-attention). Вони дають змогу враховувати глобальні зв'язки між різними частинами зображення, на відміну від CNN, які фокусуються на локальних фрагментах [4].

Сьогодні активно розвиваються гібридні архітектури, що поєднують згорткові шари для вилучення локальних ознак і блоки трансформера для аналізу глобального контексту. До того ж значного поширення набули моделі з попереднім навчанням (pre-training) на великих наборах даних, як-от ImageNet, із подальшим тонким налаштуванням (fine-tuning) під конкретні задачі [4].

3. Алгоритми навчання та оптимізації

Процес навчання нейронних мереж включає підготовку даних, вибір функції втрат, оптимізатора та технік регуляризації. Найчастіше використовуються алгоритми оптимізації Adam, SGD із моментом, RMSProp, які забезпечують стабільну збіжність. Для уникнення перенавчання застосовуються dropout, batch normalization та data augmentation – штучне розширення набору даних шляхом обертання, масштабування чи дзеркального відображення зображень [1].

Значного поширення набули методи самонавчання (self-supervised learning), що дають змогу використовувати великі обсяги неанотованих даних. Підходи, як-от SimCLR чи BYOL, формують ефективні векторні представлення зображень без потреби у ручному маркуванні [2].

Також активно досліджуються методи напівконтрольованого навчання (semi-supervised learning), що поєднують невелику кількість розмічених прикладів із великими масивами нерозмічених. Це особливо важливо для галузей, де збір даних є дорогим або складним (наприклад, у медицині).

4. Практичне використання технологій ідентифікації

Методи штучного інтелекту для ідентифікації зображень знаходять застосування у широкому спектрі галузей.

У медицині вони використовуються для виявлення патологій на рентгенівських, КТ- чи МРТ-знімках, що дає змогу підвищити точність діагностики та пришвидшити процес обробки результатів досліджень.

У сфері безпеки технології розпізнавання облич і поведінки допомагають автоматизувати системи відеоспостереження.

У транспорті вони забезпечують автоматичне зчитування номерних знаків, виявлення пішоходів і транспортних засобів у режимі реального часу.

У промисловості методи III використовуються для контролю якості продукції, виявлення дефектів або оптимізації виробничих процесів [5].

Також активно впроваджуються рішення для агросектору, де комп'ютерний зір допомагає аналізувати стан посівів за допомогою знімків із безпілотників і супутників [5].

Такі системи сприяють зменшенню впливу людського фактора, підвищенню швидкості прийняття рішень і ефективності контролю.

5. Перспективи розвитку

Подальший розвиток методів ідентифікації зображень пов'язаний із підвищенням інтерпретованості моделей і створенням систем, здатних пояснювати прийняті рішення. Цей напрям отримав назву Explainable Artificial Intelligence (XAI) [4].

Іншою перспективною тенденцією є використання генеративних моделей (Generative Adversarial Networks, GANs) для створення синтетичних навчальних даних, що дає змогу збільшити розмір тренувальних вибірок і підвищити стійкість моделей до зовнішніх впливів [3].

Також активно розвивається напрям мультисенсорного аналізу, де дані зображень поєднуються з інформацією з інших джерел – звуку, тексту, температури чи просторових сенсорів. У майбутньому це дасть змогу створювати більш адаптивні та універсальні інтелектуальні системи.

Висновки. Застосування методів штучного інтелекту для ідентифікації зображень забезпечує якісно новий рівень автоматизації та аналітики у багатьох галузях. Використання згорткових мереж, трансформерів та гібридних моделей дає змогу досягати високої точності розпізнавання навіть у складних умовах.

Подальший розвиток цієї сфери пов'язаний із підвищенням прозорості роботи моделей, розширенням доступу до великих даних і інтеграцією ШІ у повсякденні технологічні процеси. Отже, ідентифікація зображень на основі штучного інтелекту стає фундаментом для створення нових поколінь інтелектуальних систем, здатних самостійно аналізувати візуальну інформацію та приймати рішення.

Список використаних джерел

1. LeCun Y., Bengio Y., Hinton G. Deep learning. *Nature*. 2015. Vol. 521. P. 436–444.
2. Krizhevsky A., Sutskever I., Hinton G. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*. 2012. Vol. 25. P. 1097–1105. URL: https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf
3. Deep Residual Learning for Image Recognition / K. He, X. Zhang, S. Ren, J. Sun. *arXiv:1512.03385*. 2016.
4. An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale / A. Dosovitskiy et al. *arXiv:2010.11929*. 2020. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0168169917308803?via%3Dihub>
5. Kamilaris A., Prenafeta-Boldú F. X. Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*. 2018. Vol. 147. P. 70–90.