

ВИКОРИСТАННЯ МОДИФІКОВАНОЇ НЕЙРОННОЇ МЕРЕЖІ ДЛЯ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ

Вінницький національний технічний університет

Анотація. У статті розглянуто застосування модифікованої нейронної мережі для розпізнавання зображень. Проаналізовано методи побудови існуючих комп'ютерних нейромережових систем розпізнавання зображень. Описано, як саме адаптовані нейронні мережі обробляють візуальні дані, щоб ідентифікувати та класифікувати різні об'єкти на фотографіях та відео. Основна увага буде приділятися принципам функціонування таких мереж та їхній здатності до виділення ключових ознак зображень для досягнення високої точності розпізнавання.

Ключові слова: нейронна мережа, розпізнавання зображень, машинне навчання, медіа-об'єкт.

Abstract. This article examines the application of a modified neural network for image recognition. It analyzes methods for constructing existing computer neural network-based image recognition systems. The paper describes how adapted neural networks process visual data to identify and classify various objects in photographs and videos. The main focus will be on the principles of operation of such networks and their ability to extract key image features to achieve high recognition accuracy.

Keywords: neural network, image recognition, machine learning, media object.

Вступ

Використання модифікованої нейронної мережі для розпізнавання зображень відкриває широкі можливості для автоматизації та інтелектуального аналізу візуальних даних у різних сферах – від промисловості до аграрного сектору. Завдяки глибокому навчанню та адаптації під конкретні завдання, такі моделі здатні з високою точністю ідентифікувати об'єкти, їхні стани, просторові розташування або навіть приховані дефекти. Наприклад, у виробництві це дозволяє здійснювати автоматичний контроль якості продукції на конвеєрі в реальному часі – нейронна мережа може миттєво виявити тріщини, пошкодження або відхилення від стандартної форми, що значно підвищує ефективність та знижує ймовірність браку [2].

У сільському господарстві модифіковані нейромережі застосовуються для аналізу зображень рослин із дронів чи камер, щоб виявляти шкідників, хвороби або дефіцит поживних речовин. Це дає змогу аграріям оперативно реагувати на проблеми ще до того, як вони спричинять значні втрати врожаю. Крім того, подібні системи можуть працювати автономно, адаптуючись до умов конкретного середовища – змін освітлення, структури фону, типу об'єктів тощо. Такий підхід забезпечує гнучкість і масштабованість у використанні, роблячи штучний інтелект на основі зображень не лише потужним інструментом аналізу, а й надійним помічником у прийнятті рішень.

Отже, використання нейронної мережі для розпізнавання зображень стає невід'ємною частиною багатьох галузей, де потрібен глибокий та швидкий аналіз візуальної інформації[1].

Принципи роботи нейронних мереж для розпізнавання зображень

Нейронні мережі для розпізнавання зображень – це складні математичні моделі, натхненні роботою людського мозку, які здатні автоматично аналізувати й класифікувати візуальну інформацію. В основі їхньої роботи лежить принцип багатошарового обчислення: вхідне зображення проходить через кілька рівнів обробки, де на кожному етапі мережа вивчає дедалі складніші ознаки – від простих ліній і контурів до складних структур, як-от обличчя, об'єкти або сцени.

Процес починається з попередньої обробки зображення, яке перетворюється на числову матрицю – пікселі зберігаються у вигляді чисел, що відповідають яскравості або кольору. Ця матриця подається на вхід штучної нейронної мережі, найчастіше – згорткової нейронної мережі (Convolutional Neural Network, CNN), яка спеціально оптимізована для роботи з просторовими даними. Кожен нейрон у мережі отримує частину інформації, обробляє її через математичні операції (наприклад, згортки, активації) та передає далі по мережі.

На ранніх шарах CNN мережа вивчає базові візерунки – краї, кути, текстури. Ці низькорівневі ознаки комбінуються на середніх шарах у складніші структури – форми, фрагменти об'єктів. На

фінальних шарах мережа вже оперує високорівневими уявленнями, наприклад, розпізнає обличчя, транспорт, тварин або літери. В результаті цього багаторівневого аналізу формується вектор ознак – стислий числовий опис зображення, який передається в класифікатор [3].

Класифікатор – це останній етап обробки, де нейромережа «приймає рішення» про те, що саме зображено. Зазвичай він складається з одного або кількох повнозв'язних шарів, які обчислюють ймовірності належності зображення до кожної з можливих категорій. Результатом є список класів (наприклад, «кіт», «собака», «автомобіль») із відповідними ймовірностями, де найвища з них вважається остаточним розпізнаванням.

Щоб нейронна мережа могла ефективно розпізнавати зображення, вона потребує навчання на великій кількості прикладів. У процесі навчання мережа «вчиться» знаходити правильні шаблони, мінімізуючи помилки шляхом коригування вагових коефіцієнтів між нейронами. Цей процес називається зворотним поширенням помилки (backpropagation) і реалізується за допомогою оптимізаційних алгоритмів, таких як градієнтний спуск.

Таким чином, принцип роботи нейронних мереж для розпізнавання зображень базується на ієрархічному аналізі вхідних даних від елементарних візуальних ознак до складних об'єктів та сцен. Їхня здатність до самонавчання, гнучкість і точність роблять ці системи надзвичайно ефективними для багатьох практичних застосувань – від медичної діагностики до безпілотного транспорту й автоматизації виробництва [4].

Технології модифікованої нейронної мережі для розпізнавання зображень

Модифіковані нейронні мережі для розпізнавання зображень – це вдосконалені архітектури штучного інтелекту, які адаптовані під конкретні завдання аналізу візуальних даних. Вони є логічним продовженням класичних згорткових нейронних мереж (CNN), проте містять додаткові шари, механізми або стратегії обробки, що підвищують точність, швидкість або здатність працювати з обмеженими ресурсами. Основна мета таких модифікацій – зробити моделі більш ефективними, універсальними й придатними для реального використання в умовах складного середовища.

Однією з ключових технологій, що використовується в модифікованих нейронних мережах, є механізм передтренування на великих наборах даних, з подальшим донавчанням (fine-tuning) на конкретному завданні. Це дозволяє зменшити час і обсяг даних, необхідних для навчання моделі, водночас зберігаючи її високу ефективність. Завдяки цьому навіть невеликі проекти можуть отримати доступ до потужних моделей без потреби у створенні їх з нуля.

Ще одна важлива технологія – трансформерні архітектури, зокрема Vision Transformer (ViT). Вони вперше були використані у сфері обробки природної мови, проте їх модифікації продемонстрували вражаючі результати й у задачах розпізнавання зображень. Такі моделі аналізують зображення не через згортки, а через блоки уваги (attention), що дозволяє краще враховувати контекст і просторові взаємозв'язки між елементами зображення [5].

Також активно використовуються технології глибокої оптимізації, зокрема:

- 1) нормалізація (BatchNorm, LayerNorm);
- 2) регуляризація (Dropout);
- 3) розширення даних (data augmentation);
- 4) використання спеціальних функцій активації (наприклад, ReLU6, Swish, Mish).

Ці методи дозволяють уникати перенавчання, зменшити втрати точності та підвищити стійкість моделі до шуму або спотворень.

Важливим напрямом є спрощення та прискорення моделей для мобільних пристроїв і вбудованих систем – так звані lightweight-моделі. Найпоширенішими прикладами є MobileNet, EfficientNet та SqueezeNet. Такі архітектури мають компактну структуру, низьке енергоспоживання та можуть працювати без доступу до серверів, що робить їх зручними для використання в смартфонах, дронах або роботах [6].

Не менш значущими є технології пояснюваного ШІ (Explainable AI), які дозволяють візуалізувати, як саме мережа «приймає рішення», які ділянки зображення вона вважає найбільш важливими. Методи Grad-CAM, LIME або attention heatmaps допомагають інтерпретувати результати моделі, що особливо цінно в критичних галузях – медицині, судовій експертизі чи безпеці.

У підсумку, технології модифікованих нейронних мереж – це поєднання різноманітних архітектур, алгоритмів оптимізації, стратегій тренування та інтерпретації, що разом створюють потужні, адаптивні та надійні системи розпізнавання зображень. Їх розвиток не лише підвищує якість комп'ютерного зору,

а й робить його доступним у дедалі ширшому спектрі застосувань: від побутових гаджетів до складних наукових досліджень [7].

Галузі застосування систем розпізнавання зображень

Системи розпізнавання зображень на основі штучного інтелекту знайшли широке застосування в багатьох галузях, де потрібен точний і швидкий аналіз візуальної інформації. Однією з провідних сфер є медицина, де такі системи використовуються для діагностики захворювань на основі медичних зображень – зокрема, рентгенівських знімків, КТ, МРТ чи дерматологічних фото. Нейронні мережі допомагають виявляти патології, як-от пухлини, тромби або шкірні захворювання, з точністю, що подекуди перевищує людські можливості, і дозволяють проводити попередню діагностику без втручання лікаря.

У промисловості розпізнавання зображень застосовується для автоматичного контролю якості продукції. Камери та нейромережі аналізують вироби на конвеєрах у реальному часі, виявляючи дефекти, тріщини, подряпини або відхилення від стандартних форм. Це значно підвищує ефективність виробництва, зменшує втрати та людський фактор. Крім того, такі системи використовуються для моніторингу процесів, наприклад, у хімічному або харчовому виробництві.

У сільському господарстві системи аналізу зображень допомагають визначати стан рослин, виявляти шкідників, хвороби чи нестачу поживних речовин. Зйомка полів з дронів або супутників у поєднанні з алгоритмами розпізнавання дозволяє аграріям оперативно реагувати на проблеми, оптимізувати зрошення та добриво, що підвищує врожайність та економить ресурси. Також ці технології використовуються в автоматизованому збиранні врожаю та сортуванні продуктів.

Ще однією перспективною галуззю є роздрібна торгівля та логістика. Системи розпізнавання зображень використовуються для моніторингу полиць у магазинах, контролю залишків товару, аналізу поведінки покупців, а також для ідентифікації товарів на складах. У поєднанні з автоматизованими системами обліку це дозволяє покращити управління ланцюгами постачання, зменшити витрати та підвищити якість обслуговування. Розпізнавання зображень також активно застосовується в системах безпеки, доповненій реальності, робототехніці та багатьох інших сучасних технологіях [8].

Висновки

Впровадження модифікованої нейронної мережі для розпізнавання зображень є важливим кроком у напрямку автоматизації аналізу візуальної інформації в найрізноманітніших галузях – від медицини до промисловості, сільського господарства та логістики. Завдяки сучасним технологіям, зокрема оптимізованим моделям для мобільних пристроїв і методам пояснюваного штучного інтелекту, такі мережі забезпечують високу точність, адаптивність і швидкість обробки зображень. Їх здатність самонавчатися, працювати в режимі реального часу та ефективно використовувати ресурси дає змогу розв'язувати складні задачі в умовах обмежених даних або змінного середовища. Особливо цінним є застосування модифікованих нейромереж там, де людський контроль є складним або затратним – наприклад, під час діагностики захворювань, моніторингу стану врожаю чи контролю якості на виробництві. Такі системи не лише покращують точність і продуктивність, а й створюють нові підходи до управління, аналізу та прийняття рішень. Загалом, розвиток і впровадження подібних технологій змінює сучасні уявлення про обробку зображень, перетворюючи штучний інтелект на незамінного помічника в реальних практичних задачах.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Kamilaris, A., & Prenafeta-Boldú, F. X. Deep learning in agriculture // *Computers and Electronics in Agriculture*, 2018, 147, 70-90.
2. Russakovsky, O. et al. ImageNet Large Scale Visual Recognition Challenge. // *IJC*, 2015, 115, 211-252.
3. Deep Learning [Електронний ресурс] – Режим доступу до ресурсу: <https://www.deeplearningbook.org/>
4. Litjens, G. et al. . A survey on deep learning in medical image analysis // *Medical Image Analysis*, 2017, 42, 60-88.
5. Transformers and Pretrained Language Models [Електронний ресурс] – Режим доступу до ресурсу: <https://av.tib.eu/media/63458>
6. Lightweight Models for Real-Time Steganalysis: A Comparison of MobileNet, ShuffleNet, and EfficientNet [Електронний ресурс] – Режим доступу до ресурсу: https://www.researchgate.net/publication/387956734_Lightweight_Models_for_Real-Time_Steganalysis_A_Comparison_of_MobileNet_ShuffleNet_and_EfficientNet
7. Selvaraju, R. R. et al. Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization // *ICCV*, 2019, 128, 336-359.

8. What is AI Image Recognition? How it works and examples [Електронний ресурс] – Режим доступу до ресурсу: <https://uk.shaip.com/blog/what-is-ai-image-recognition-and-how-does-it-work/>

Дерев'янчук Андрій Юрійович – студент групи 4ПІ-21б, ФІТКІ, Вінницький національний технічний університет, Вінниця, e-mail: andrewderev@gmail.com

Черноволік Галина Олександрівна – к.т.н., доцент, Вінницький національний технічний університет, м. Вінниця, e-mail: lina2433@gmail.com

Andrii Yuriyovych Derevianchuk – student of group 4PI-21b, FITKI, Vinnytsia National Technical University, Vinnytsia, e-mail: andrewderev@gmail.com.

Chernovolyyk Halyna Oleksandrivna – PhD, Associate Professor, Vinnytsia National Technical University, Vinnytsia, e-mail: lina2433@gmail.com.