

УДК 004.8

DOI <https://doi.org/10.32782/IT/2024-2-10>

Владислав КРОМКАЧ

аспірант кафедри комп'ютеризованих систем автоматики, Національний університет «Львівська політехніка», вул. С. Бандери, 12, м. Львів, Україна, 79013

ORCID: 0009-0001-5608-5715

Бібліографічний опис статті: Кромкач, В. (2024). Роль комп'ютерного зору в сучасному світі: досягнення, виклики та перспективи. *Information Technology: Computer Science, Software Engineering and Cyber Security*, 2, 79–87, doi: <https://doi.org/10.32782/IT/2024-2-10>

РОЛЬ КОМП'ЮТЕРНОГО ЗОРУ В СУЧАСНОМУ СВІТІ: ДОСЯГНЕННЯ, ВИКЛИКИ ТА ПЕРСПЕКТИВИ

Ця наукова стаття розглядає сутність комп'ютерного зору як сучасної технології, її досягнення, виклики та перспективи. Комп'ютерний зір відіграє важливу роль у різних галузях, таких як медицина, автомобільна промисловість, робототехніка та багато інших. У статті досліджуються основні принципи комп'ютерного зору, його застосування в різних сферах, порівняльний аналіз нейронних мереж.

Метою дослідження є вибір топології згорткових нейронних мереж шляхом проведення порівняльного аналізу моделей згорткових нейронних мереж та їх характеристик, які суттєво впливають на реалізацію типових задач комп'ютерного зору.

Методологія дослідження полягає у порівняльному аналізі основних методів класифікації на основі відомих згорткових нейронних мереж, які орієнтовані на обробку та розпізнавання образів зображень, сегментацію текстових даних, а також звукових потоків. Проведено аналіз існуючих досліджень та публікацій щодо технологій комп'ютерного зору, емпіричне дослідження та оцінку ефективності і точності різних підходів до комп'ютерного зору.

Наукова новизна отриманих у роботі результатів полягає у визначенні основних викликів, що стоять перед розвитком комп'ютерного зору, зокрема, проблеми обробки великих обсягів даних, забезпечення точності та швидкості алгоритмів, проведенні комплексного аналізу сучасних досягнень у галузі комп'ютерного зору з урахуванням останніх наукових досліджень та технологічних розробок.

Висновки. Нейромережі, такі як штучні нейронні мережі (ANN), згорткові нейронні мережі (CNN) та рекурентні нейронні мережі (RNN), мають власні переваги та недоліки, що можуть залежати від конкретних завдань та даних. Враховуючи це, доцільно їх використовувати в залежності від проблем чи завдань, які необхідно вирішити, а саме: ANN (штучні нейронні мережі) корисні для вирішення складних проблем; CNN (згорткові нейронні мережі) є найефективнішим способом вирішення проблем комп'ютерного зору; RNN (рекурентні нейронні мережі) здатні обробляти природну мову, корисні для прогнозування часових рядів.

Ключові слова: штучний інтелект, комп'ютерний зір, нейронні мережі, штучні нейронні мережі, згорткові нейронні мережі, рекурентні нейронні мережі, розпізнавання образів.

Vladyslav KROMKACH

Postgraduate student of the Computerized Automation Systems Department, Lviv Polytechnic National University, 12, S. Bandery Str., Lviv, Ukraine, 79013, vlad.kromkach@gmail.com

ORCID: 0009-0001-5608-5715

To cite this article: Kromkach, V. (2024). Rol kompiuternoho zoru v suchasnomu sviti: dosiahnennia, vyklyky ta perspektyvy [The role of computer vision in the modern world: achievements, challenges and perspectives]. *Information Technology: Computer Science, Software Engineering and Cyber Security*, 2, 79–87, doi: <https://doi.org/10.32782/IT/2024-2-10>

THE ROLE OF COMPUTER VISION IN THE MODERN WORLD: ACHIEVEMENTS, CHALLENGES AND PROSPECTS

This scientific article examines the essence of computer vision as a modern technology, its achievements, challenges and prospects. Computer vision plays an important role in various fields such as medicine, automotive, robotics and many others. The article examines the basic principles of computer vision, its application in various fields, comparative analysis of neural networks.

The purpose of the study is to choose the topology of convolutional neural networks by conducting a comparative analysis of models of convolutional neural networks and their characteristics, which significantly affect the implementation of typical computer vision tasks.

The methodology consists in a comparative analysis of the main classification methods based on well-known convolutional neural networks, which are focused on the processing and recognition of image patterns, segmentation of text data, as well as sound streams. An analysis of existing research and publications on computer vision technologies, empirical research and evaluation of the effectiveness and accuracy of various approaches to computer vision has been carried out.

The scientific novelty of the results obtained in the work consists in determining the main challenges facing the development of computer vision, in particular, the problem of processing large volumes of data, ensuring the accuracy and speed of algorithms, conducting a comprehensive analysis of modern achievements in the field of computer vision, taking into account the latest scientific research and technological development.

Conclusions. *Neural networks such as Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) have their own advantages and disadvantages that may depend on specific tasks and data. Given this, it is advisable to use them depending on the problems or tasks that need to be solved, namely: ANNs (artificial neural networks) are useful for solving complex problems; CNNs (Convolutional Neural Networks) are the most effective way to solve computer vision problems; RNNs (recurrent neural networks) are capable of natural language processing, useful for time series forecasting.*

Key words: *artificial intelligence, computer vision, neural networks, artificial neural networks, convolutional neural networks, recurrent neural networks, pattern recognition.*

Актуальність проблеми. Останнім часом зростає значення комп'ютерного зору як інноваційної технології. Комп'ютерний зір дозволяє комп'ютерам розуміти та аналізувати зображення і відео, надаючи їм здатність «бачити» світ. Машинний зір виконує такі ж самі завдання, як і людський мозок за допомогою очей. Комп'ютерне бачення використовує нейронні мережі, які обробляють зображення та відео. За останнє десятиліття досягнення в області глибинного структурного навчання і штучного інтелекту дозволили збільшити точність визначення і класифікації об'єктів з 50 до 99%.

Використання комп'ютерного зору дозволяє розпізнавати обличчя для систем безпеки. Поширення різних типів камер робить зйомку простою, дешевою та ненав'язливою. Крім того, це біометричний атрибут, який полегшує візуальне підтвердження людиною результату, отриманого технологічним рішенням, що неможливо з іншими біометричними атрибутами. Наприклад, людина не може підтвердити розпізнавання, зроблене за допомогою відбитка пальця або райдужної оболонки ока, але це набагато легше зробити за допомогою обличчя (Marta Beltrán, Miguel Calvo, 2023).

Комп'ютерний зір необхідний для створення самокерованих автомобілів, даючи можливість виявляти об'єкти, розмітку смуг руху, знаки та сигнали світлофора для безпечного керування автомобілем. Щоб досягти виявлення об'єктів, пізнання та сприйняття сцени, очікується, що безпілотні автомобілі сприйматимуть оточення у спосіб, щонайменше подібний до того, як людське око обробляє інформацію, що веде до когнітивних систем штучного інтелекту, які можуть вчитися, перенавчатися, виконувати дії. Безпілотні автомобілі повинні вміти розпізнавати навколишнє середовище, інтерпретувати 3D-репрезентацію світу, розпізнавати рух

об'єктів, пішоходів та інших автомобілів і мати справу з людськими емоціями (Abhishek Gupta, Alagan Anpalagan, Ling Guan, Ahmed Shaharyar Khwaja, 2021).

Комп'ютерне бачення дає можливість здійснювати автоматичний контроль у сучасних роздрібних магазинах. Автоматизовані роздрібні магазини на базі штучного інтелекту (AIPARS) можна визначити як магазини, які інтегрують штучний інтелект і робототехніку разом із передовими системами програмного забезпечення для створення повністю автоматизованих роздрібних магазинів самообслуговування, які надають покупцям оплату в один клік, автоматизовані пояснення, маркетинг, продажі, платежі, налаштування сповіщень про продукт, інформацію про доставку продукту та навчання клієнтів (Rajasshrie Pillai, Brijesh Sivathanu, Yogesh K. Dwivedi, 2020).

У медицині комп'ютерний зір можна використовувати для виявлення ознак захворювання на зображеннях МРТ або рентгенівських знімках. З моменту свого розквіту глибоке навчання продемонструвало надзвичайний успіх у автоматичному аналізі зображень методів медичної візуалізації. Прогрес у глибокому навчанні процвітає і вдосконалювався з часом, головним чином обертаючись навколо одного алгоритму, який називається згортковими нейронними мережами (CNN). CNN є потенційно найпопулярнішою архітектурою глибокого навчання завдяки своїм видатним можливостям використовувати просторовий і часовий зв'язок між характеристиками зображень, які необхідно розшифрувати для вилучення значущої інформації, прихованої в зображеннях. Він досяг значних досягнень у застосуванні медичної візуалізації, таких як визначення наявності та визначення типу злоякісного новоутворення (класифікація), визначення місця ураження пацієнта (виявлення),

виділення потрібного об'єкта (органу) з медичного зображення (сегментація), розміщення окремих зображень у загальній системі відліку для порівняння або інтеграції інформації, яку вони містять (реєстрація), синтезування зображень для збалансування набору даних (генеративне моделювання) (4).

Комп'ютерне бачення забезпечує роботу автоматичних паркувальних майданчиків, дає можливість контролювати рух транспортних засобів. Система запобігання зіткненням відіграє важливу роль у зниженні кількості інцидентів на дорозі, причому виявлення об'єктів має вирішальне значення для уникнення перешкод для цієї системи (Jittima Varagula, Panit-a-nong Kulproma, Toshio ITO, 2017).

Сучасні аеропорти часто мають великі та складні зони злітно-посадкової зони з кількома злітно-посадковими смугами зі змінною конфігурацією, численними доріжками для ефективної циркуляції рейсів і десятками, якщо не сотнями виходів. Ефективне спостереження та управління операціями в зоні злітно-посадкової зони є надзвичайно складним завданням для авіадиспетчерів. Використання згорткових нейронних мереж для спостереження за зоною злітно-посадкової зони з використанням камер для моніторингу об'єктів, що рухаються на землі, дасть можливість підвищити безпеку та ефективність роботи (Phat Thai, Sameer Alam, Nimrod Lilith, Binh T. Nguyen, 2022).

Технології комп'ютерного зору в поєднанні з отриманням високоякісних зображень за допомогою дистанційних камер дозволяють використовувати безконтактні та ефективні технологічні рішення в сільському господарстві. Основним застосуванням глибокого навчання в галузі сільського господарства є створення моделей для отримання значущої інформації з даних про сільське господарство, аналіз зображень, включаючи класифікацію та виявлення об'єктів, наприклад виявлення хвороби, ідентифікація бур'янів, аналіз ґрунту, виявлення хвороб рослин тощо (7).

Виробниче обладнання контролюється за допомогою комп'ютерного зору, щоб втрутитися до того, як поломка спричинить дороге простоювання. Для виявлення зносу оцінюються моделі класифікації та регресії з використанням згорткових нейронних мереж (8).

Комп'ютерне бачення допомагає в реальному часі аналізувати спортивну гру та стратегію, продуктивність і рейтинги гравців. Розпізнавання дій людини – це складне завдання, яке використовується в таких видах спорту, як волейбол, баскетбол, футбол і теніс, для

виявлення гравців і розпізнавання їхніх дій і дій команд під час тренувань, матчів, розминки або змагань (Phat Thai, Sameer Alam, Nimrod Lilith, Binh T. Nguyen, 2022).

Аналіз останніх досліджень і публікацій.

Комп'ютерний зір використовує алгоритми, які дають можливість комп'ютеру обробляти зображення та відео, аналізувати їх і виокремлювати з них корисну інформацію, обробляючи великі масиви даних.

Основні алгоритми, які застосовуються в комп'ютерному зорі, – це алгоритми виокремлення ознаки та алгоритми класифікації. Алгоритми виокремлення ознаки для виявлення ключових рис зображень, таких як грані, кути та кольорові плями, дозволяють комп'ютеру розпізнати об'єкти на зображеннях. Алгоритми класифікації використовують для визначення класу об'єкта на зображенні.

Серед недоліків існуючих алгоритмів, які застосовуються в комп'ютерному зорі, для обробки зображень можна виокремити наступні:

- при розпізнаванні об'єкта один чи декілька об'єктів можуть бути розпізнані разом із їх двовимірним положенням на зображенні/тривимірним положенням на сцені;
- точність ідентифікації при розпізнаванні та порівнянні індивідуальних біометричних даних об'єкта;
- правильна інтерпретація при виявленні (наприклад, виявлення можливих змін клітин/тканин та їх властивостей при медичних обстеженнях, перешкод в системах керування робототехнікою, тощо).

Методи глибокого навчання зазвичай вимагають величезного набору даних, щоб мати можливість успішно узагальнювати нові екземпляри даних і фіксувати дрібні деталі в даних. Таким чином, неможливо навчити методів глибокого навчання для розпізнавання людської діяльності з невеликими наборами даних, оскільки недостатньо навчальних даних, особливо для незвичайних дій (A. Nouriani, R. McGovern, R. Rajamani, 2023).

Оскільки збільшення швидкодії обробки даних призводить до зниження якості, актуальними напрямками досліджень є підвищення швидкості обробки існуючих алгоритмів; застосування алгоритмів розпізнавання в уже реалізованих автоматичних системах; вирішення задач в режимі реального часу в умовах обмежених ресурсів; розпізнавання зображень нейронними мережами з підвищеною точністю, якістю та розширеним спектром.

Мета і задачі дослідження. Метою дослідження є вибір топології згорткових нейронних

мереж шляхом проведення порівняльного аналізу моделей згорткових нейронних мереж та їх характеристик, які суттєво впливають на реалізацію типових задач комп'ютерного зору.

Основною задачею дослідження є огляд методів класифікації на основі відомих згорткових нейронних мереж, які орієнтовані на обробку та розпізнавання образів зображень, сегментації текстових даних, а також звукових потоків.

Виклад основного матеріалу.

Порівняльний огляд нейронних мереж

Одним із найпоширеніших методів, які використовуються у комп'ютерному зорі, є нейронні мережі. Ці мережі складаються з багатьох шарів нейронів, які разом працюють для виявлення та класифікації об'єктів на зображеннях. Вони можуть автоматично розпізнавати шаблони та ознаки на зображеннях, що робить їх особливо ефективними для завдань комп'ютерного зору.

Штучні нейронні мережі (ANN)

Завдяки своїй багаторівневій багатозаровності штучні нейронні мережі (ANN) використовуються в технології, зосередженій на розв'язанні складних проблем, таких як проблеми розпізнавання образів. Їх принцип роботи схожий до роботи нейронів у нашому мозку: розпізнаючи шаблони в необроблених даних, вирішуються складні процеси, а результати покращуються з кожним новим вхідним сигналом.

Штучна нейронна мережа базується на трьох або більше взаємопов'язаних шарах вузлів. Усі рівні відповідають за введення, обробку та виведення даних на більш глибокі рівні. Саме така багаторівнева система допомагає цій нейронній мережі розуміти та вивчати складні речі.

Використання ANN може принести багато переваг, ось деякі з основних:

- адаптивний характер: нейронна мережа модифікує себе після початкового навчання з подальшою обробкою інформації;

- покращена здатність до навчання: ANN може вивчати події самостійно, може приймати рішення або генерувати результати, порівнюючи проблему з подібними подіями, завдяки потужності вузла ANN можуть виконувати декілька завдань одночасно;

- розподілене зберігання інформації: ANN зберігає інформацію про всю мережу, а не про конкретну базу даних, кілька фрагментів відсутньої інформації в одному місці не виводять із ладу всю систему.

Однак, незважаючи на численні переваги, штучні нейронні мережі можуть спричинити деякі труднощі:

- для навчання ANN потрібна велика кількість даних;

- складність навчання мережі;

- архітектура може бути нечіткою. Коли ANN дає рішення проблеми, воно чітко не вказує, чому і як було досягнуто цього результату (11).

Згорткові нейронні мережі (CNN)

Згорткова нейронна мережа (CNN) – це особливий вид нейронної мережі, який значно зменшує кількість параметрів у глибокій нейронній мережі з багатьма одиницями, не надто втрачаючи якість моделі. CNN знайшли застосування в обробці зображень і тексту, розпізнаванні зображень, класифікації зображень, виявленні об'єктів, розпізнаванні облич тощо.

Для ідентифікації предметів на зображеннях раніше використовувалася ручна праця. Однак CNN допоміг масштабувати процес,

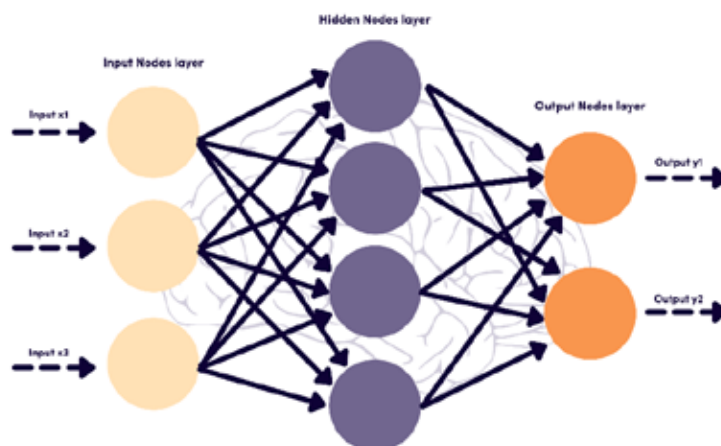


Рис. 1. Як працюють штучні нейронні системи

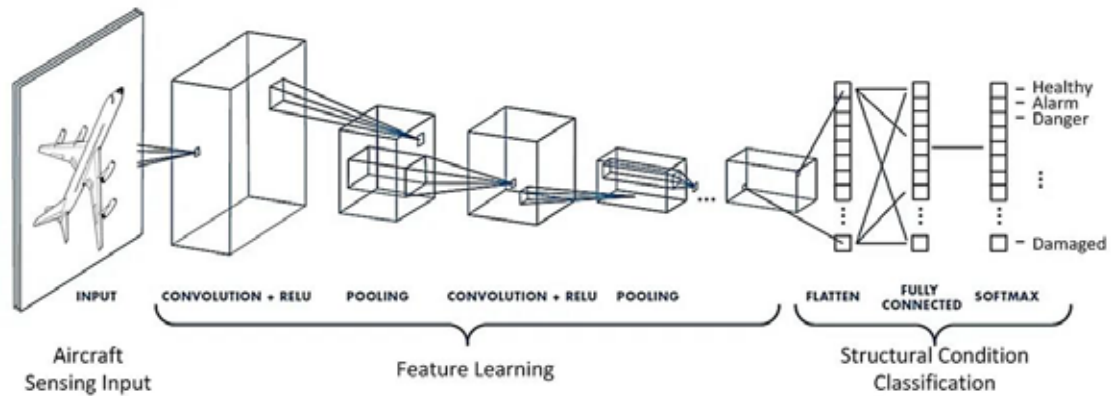


Рис. 2. Загальна архітектура згорткової нейронної мережі CNN (12).

використовуючи принципи лінійної алгебри для виявлення шаблонів на зображеннях.

У структурі згорткової нейронної мережі основними видами прошарків є:

- згорткові прошарки (convolutional layer);
- пулінгові прошарки (pooling layer);
- повнозв'язані прошарки (fully-connected layer).

З кожним шаром CNN ускладнює розуміння зображення. Це означає, що перші шари зосереджені на інтерпретації простих елементів зображення, таких як його краї та кольори. Коли зображення обробляється через шари, мережа здатна розпізнавати складні особливості, такі як форми об'єктів. Нарешті, найглибший рівень здатний ідентифікувати цільовий об'єкт.

Типовим випадком використання CNN є те, що ви подаєте образи мережі, і мережа класифікує дані. CNN, як правило, починаються з вхідного «сканера», який не призначений для аналізу всіх навчальних даних одночасно. Наприклад, щоб ввести зображення розміром 100 x 100 пікселів, вам не потрібен шар із 10000 вузлів (13).

Згорткові нейронні мережі, загалом, використовуються для комп'ютерного зору. Головна перевага CNN – автоматично визначає функції без нагляду людини, що дає їй перевагу над попередниками.

Проблеми використання CNN:

- труднощі в роботі з розбіжностями представлених даних. CNN важко обробляти об'єкти на зображеннях, які певною мірою приховані. Мережа також має труднощі з класифікацією заголовків або повернутих зображень. Простіше кажучи, CNN не може кодувати орієнтацію та положення об'єкта та не може обробляти просторово інваріантні дані.

- вимоглива до обчислень. Для навчання CNN потрібні численні графічні процесори

(GPU). До того ж, якщо не вистачає хороших графічних процесорів, навчання стає повільним.

Існуючі ітераційні методи скорочення для CNN не мають ефективних стратегій навчання для відновлення точності стисненої мережі між кожними двома кроками скорочення. Коли втрата продуктивності безперервно зростає зі стисненням, виникає проблема неточної оцінки важливості параметра, що призводить до низької продуктивності кінцевої отриманої стисненої мережі (14).

Порівнюючи переваги та недоліки використання нейромереж ANN та CNN, можна зробити висновок, що існують конкретні випадки, коли ANN можна надати перевагу над CNN і навпаки. Вони обидві унікальні в тому, як вони працюють математично, і це змушує їх краще вирішувати конкретні проблеми. Загалом CNN є більш потужним і точним способом вирішення проблем класифікації. ANN все ще домінує для проблем, де набори даних обмежені, а вхідні зображення не потрібні. Однак завдяки здатності CNN переглядати зображення як дані, це найпоширеніше рішення для проблем комп'ютерного зору та машинного навчання, що залежить від зображень (15).

Рекурентні нейронні мережі (RNN)

RNN є унікальними завдяки своїй здатності обробляти як минулі дані, так і вхідні дані та запам'ятовувати. Вони були розроблені для подолання слабких місць прямої мережі. Практичні програми включають голосовий пошук від Google і Siri від Apple.

Подібно до ANN і CNN, нейромережа RNN також навчається за допомогою навчальних даних, вона не обробляє лише введені дані, але також використовує минулі вхідні дані для прийняття рішень.

У традиційній нейронній мережі входи та виходи не залежать один від одного, тоді як

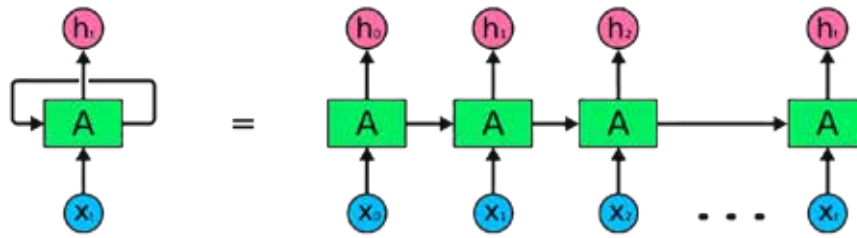


Рис. 3. Рекурентні нейронні мережі RNN

вихід у RNN залежить від попередніх елементів у послідовності. Рекурентні мережі також спільно використовують параметри на кожному рівні мережі. У мережах прямого зв'язку для кожного вузла існують різні ваги. У той час як RNN має однакові ваги на кожному рівні мережі та під час градієнтного спуску, ваги та базис налаштовуються окремо, щоб зменшити втрати (16).

На відміну від CNN, які мають прямі з'єднання, рекурентні нейронні мережі (RNN) мають зв'язки, які передають активації від входу на попередньому часовому етапі назад у мережу, щоб впливати на вихід для поточного входу (Kristina Host, Marina Ivašić-Kos, 2022).

Оскільки RNN має внутрішню пам'ять, вона може робити відносно точні прогнози. Більше того, це корисно для вирішення послідовних проблем на основі даних. Рекурентні нейронні мережі мають різні мережеві структури. Загальні структури включають один до одного, один до багатьох, багато до одного та багато до багатьох.

Рекурентні нейронні мережі мають унікальні можливості на відміну від інших типів нейронних мереж, які відкривають широкі можливості для своїх користувачів, але також приносять із собою певні труднощі. Це єдина нейронна мережа з пам'яттю та подвійною обробкою даних.

Ось короткий перелік основних переваг:

- RNN можуть маніпулювати даними послідовності;
- RNN можуть обробляти вхідні дані різної довжини;
- RNN мають пам'ять, точніше можуть зберігати історичні дані.

Наступний список містить недоліки рекурентної нейронної мережі:

- обчислення можуть бути повільними;
- RNN не можуть враховувати майбутні вхідні дані для прийняття рішень;
- градієнти, які використовуються для обчислення оновлених ваг, можуть бути дуже

близькими до нуля, заважаючи мережі вивчати нові ваги. Це можна назвати проблемою зникаючого градієнта;

- під час навчання мережі дані можуть експоненціально зростати, а не спадати, що називається проблемою вибухового градієнта (18).

Рекурентні мережі RNN застосовуються для машинного перекладу, розпізнавання мови, прогнозу часових рядів, аналізу настроїв, прогнозування курсу акцій, генерації тексту та моделювання мови, вирішення проблеми прогнозування, тощо.

Кожна з розглянутих нейронних мереж має свої переваги та недоліки, які узагальнено в таблиці 1.

Висновки і перспективи подальших досліджень

Комп'ютерний зір – це сфера штучного інтелекту (ШІ), яка дозволяє комп'ютерам і системам отримувати важливу інформацію з цифрових зображень, відео та інших візуальних вхідних даних і вживати дії або давати рекомендації на основі цієї інформації. Якщо ШІ дозволяє комп'ютерам мислити, комп'ютерний зір дозволяє їм бачити, спостерігати та розуміти.

Для досягнення точних результатів у комп'ютерному баченні потрібні алгоритми машинного навчання та нейронні мережі, які можуть використовуватися для розпізнавання зображень і відео, підвищуючи точність комп'ютерного зору.

В результаті проведеного дослідження та порівняльного огляду нейронних мереж можна зробити висновок, що нейромережі, такі як штучні нейронні мережі (ANN), згорткові нейронні мережі (CNN) та рекурентні нейронні мережі (RNN), мають власні переваги та недоліки, що можуть залежати від конкретних завдань та даних. Враховуючи це, доцільно їх використовувати в залежності від проблем чи завдань, які необхідно вирішити, а саме:

- ANN (штучні нейронні мережі) корисні для вирішення складних проблем;

Таблиця 1

Порівняння характеристик CNN, ANN і RNN

	ANN	CNN	RNN
Опис	Один з найпростіших типів нейронних мереж	Один з найпопулярніших типів нейронних мереж	Найбільш передова і комплексна нейронна мережа
Структура	Її простота походить від її прямої природи – інформація тече лише в одному напрямку	Структура складається з декількох слоїв, включаючи 1 або більше згорткових слоїв	Інформація тече в різних напрямках, що надає їй власну пам'ять та можливість до самонавчання
Архітектура	Складається з великої кількості шарів, в яких нейрони пов'язані з попередніми та наступними шарами. Кожен нейрон у верхньому шарі пов'язаний з кожним нейроном у наступному шарі.	Має додаткову структуру, включаючи згорткові шари, шари підвищення деталізації (pooling) та повністю зв'язані шари. Згорткові шари можуть відокремити важливі особливості зображення.	RNN призначені для роботи з послідовними даними. Вони мають зв'язки між нейронами, які можуть передавати інформацію від попередніх кроків до наступних. Використовуються для аналізу наслідків, таких як мовні дані або часові ряди.
Тип даних	Базується на табличних і текстових даних	Базується на зображеннях	Тренується послідовними даними
Обробка образів	Може бути менш ефективною для обробки образів через велику кількість параметрів та відображення рахування просторової структури.	Оптимально підходить для роботи з обробкою зображення за допомогою спеціалізованої архітектури, яка дозволяє виявляти локальні особливості та ієрархічні структури в зображеннях.	Зазвичай рідше використовується для обробки зображення, але може використовуватися для аналізу сторінок у контексті обробки образів, наприклад, у завданнях генерації підписів до зображення.
Складність	Легка у порівнянні з 2 іншими моделями	Вважається більш потужною, ніж 2 інші методи	Менше функцій, ніж у CNN, але потужна через потенціал до самонавчання і пам'яті
Переваги	Здатність працювати з неповними даними та висока відмовостійкість	Точність розпізнавання зображення	Пам'ять і самонавчання
Просторове розпізнавання	Ні	Так	Ні
Повторювані з'єднання	Ні	Ні	Так
Основні недоліки	Залежність від апаратної складової	Потребує багато даних для навчання	Повільне і складне навчання та проблеми з градієнтом
Використання	Вирішення комплексних проблем, таких як прогнозування результату	Комп'ютерне бачення, включаючи розпізнавання зображень	Обробка природної мови, включаючи аналіз настроїв і розпізнавання швидкості

– CNN (згорткові нейронні мережі) є найефективнішим способом вирішення проблем комп'ютерного зору;

– RNN (рекурентні нейронні мережі) здатні обробляти природну мову, корисні для прогнозування часових рядів.

Отже, комп'ютерний зір – це потужний інструмент, який знайшов застосування в багатьох галузях. Розвиток цієї технології продовжується, а використання комп'ютерного зору розширюється та відкриває нові можливості для людства.

ЛІТЕРАТУРА:

1. Marta Beltrán, Miguel Calvo. A privacy threat model for identity verification based on facial recognition. 2023. URL: <https://doi.org/10.1016/j.cose.2023.103324> (дата звернення: 18.04.2024).
2. Abhishek Gupta, Alagan Anpalagan, Ling Guan, Ahmed Shaharyar Khwaja. Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues. 2021. <https://doi.org/10.1016/j.array.2021.100057> (дата звернення: 18.04.2024).
3. Rajasshrie Pillai, Brijesh Sivathanu, Yogesh K. Dwivedi. Shopping intention at AI-powered automated retail stores (AIPA). Journal of Retailing and Consumer Services. 2020. URL: <https://doi.org/10.1016/j.jretconser.2020.102207> (дата звернення: 18.04.2024).

4. Arshi Parvaiz, Muhammad Anwaar Khalid, Rukhsana Zafar, Huma Ameer, Muhammad Ali, Muhammad Moazam Fraz. Vision Transformers in medical computer vision—A contemplative retrospection. *Engineering Applications of Artificial Intelligence*. 2023. URL: <https://doi.org/10.1016/j.engappai.2023.106126> (дата звернення: 19.04.2024).
5. Jittima Varagula, Panit-a-nong Kulproma, Toshio ITO. Object Detection Method in Traffic by On-Board Computer Vision with Time Delay Neural Network. *Procedia Computer Science*. 2017. URL: <https://doi.org/10.1016/j.procs.2017.08.185> (дата звернення: 19.04.2024).
6. Phat Thai, Sameer Alam, Nimrod Lilith, Binh T. Nguyen. A computer vision framework using Convolutional Neural Networks for airport-airside surveillance. *Transportation Research Part C: Emerging Technologies*. 2022. URL: <https://doi.org/10.1016/j.trc.2022.103590> (дата звернення: 19.04.2024).
7. V.G. Dhanya, A. Subeesh, N.L. Kushwaha, Dinesh Kumar Vishwakarma, T. Nagesh Kumar, G. Ritika, A.N. Singh. Deep learning based computer vision approaches for smart agricultural applications. 2022. URL: <https://doi.org/10.1016/j.aiaa.2022.09.007> (дата звернення: 26.04.2024).
8. Markus Friedrich, Theresa Gerber, Jonas Dumler, Frank Döpfer. A system for automated tool wear monitoring and classification using computer vision. *Procedia CIRP*. 2023. URL: <https://doi.org/10.1016/j.procir.2023.06.073> (дата звернення: 26.04.2024).
9. Phat Thai, Sameer Alam, Nimrod Lilith, Binh T. Nguyen. An overview of Human Action Recognition in sports based on Computer Vision. *Journal Heliyon*. 2022. URL: <https://doi.org/10.1016/j.heliyon.2022.e09633> (дата звернення: 26.04.2024).
10. A. Nouriani, R. McGovern, R. Rajamani. Activity recognition using a combination of high gain observer and deep learning computer vision algorithms. *Intelligent Systems with Applications*. 2023. URL: <https://doi.org/10.1016/j.iswa.2023.200213> (дата звернення: 02.05.2024).
11. Masooma Memon. ANN vs CNN vs RNN: Neural Networks Guide. 2022. URL: <https://levity.ai/blog/neural-networks-cnn-ann-rnn> (дата звернення: 02.05.2024).
12. Zahra Elhamraoui, Introduction to convolutional neural network. 2020. URL: <https://medium.com/analytics-vidhya/introduction-to-convolutional-neural-network-6942c189a723> (дата звернення: 02.05.2024).
13. James Le The 5 Computer Vision Techniques That Will Change How You See The World. 2018. URL: <https://heartbeat.comet.ml/the-5-computer-vision-techniques-that-will-change-how-you-see-the-world-1ee19334354b> (дата звернення: 02.05.2024).
14. Jingfei Chang, Yang Lu, Ping Xue, Yiqun Xu, Zhen Wei. Iterative clustering pruning for convolutional neural networks. *Knowledge-Based Systems*. 2023. URL: <https://doi.org/10.1016/j.knosys.2023.110386> (дата звернення: 03.05.2024).
15. Vidushi Meel. ANN and CNN: Analyzing Differences and Similarities. URL: <https://viso.ai/deep-learning/ann-and-cnn-analyzing-differences-and-similarities/> (дата звернення: 03.05.2024).
16. Recurrent Neural Network Tutorial (RNN). 2022. URL: <https://www.datacamp.com/tutorial/tutorial-for-recurrent-neural-network> (дата звернення: 03.05.2024).
17. Kristina Host, Marina Ivašić-Kos. An overview of Human Action Recognition in sports based on Computer Vision. 2022. URL: <https://doi.org/10.1016/j.heliyon.2022.e09633> (дата звернення: 03.05.2024).
18. What is a recurrent network? URL: <https://h2o.ai/wiki/recurrent-network/> (дата звернення: 03.05.2024).

REFERENCES:

1. Marta Beltrán, Miguel Calvo. (2023). A privacy threat model for identity verification based on facial recognition. Retrieved from <https://doi.org/10.1016/j.cose.2023.103324> [in English].
2. Abhishek Gupta, Alagan Anpalagan, Ling Guan, Ahmed Shaharyar Khwaja. (2021). Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues. Retrieved from <https://doi.org/10.1016/j.array.2021.100057> [in English].
3. Rajasshrie Pillai, Brijesh Sivathanu, Yogesh K. Dwivedi. (2020). Shopping intention at AI-powered automated retail stores (AIPA). *Journal of Retailing and Consumer Services*. Retrieved from <https://doi.org/10.1016/j.jretconser.2020.102207> [in English].
4. Arshi Parvaiz, Muhammad Anwaar Khalid, Rukhsana Zafar, Huma Ameer, Muhammad Ali, Muhammad Moazam Fraz. (2023). Vision Transformers in medical computer vision – A contemplative retrospection. *Engineering Applications of Artificial Intelligence*. Retrieved from <https://doi.org/10.1016/j.engappai.2023.106126> [in English].
5. Jittima Varagula, Panit-a-nong Kulproma, Toshio ITO. (2017). Object Detection Method in Traffic by On-Board Computer Vision with Time Delay Neural Network. *Procedia Computer Science*. Retrieved from <https://doi.org/10.1016/j.procs.2017.08.185> [in English].

6. Phat Thai, Sameer Alam, Nimrod Lilith, Binh T. Nguyen. (2022). A computer vision framework using Convolutional Neural Networks for airport-airside surveillance. *Transportation Research Part C: Emerging Technologies*. Retrieved from <https://doi.org/10.1016/j.trc.2022.103590> [in English].
7. Dhanya, V. G., Subeesh, A., Kushwaha, N. L., Dinesh Kumar Vishwakarma, Nagesh Kumar, T., Ritika, G., Singh. A. N. (2022). Deep learning based computer vision approaches for smart agricultural applications. Retrieved from <https://doi.org/10.1016/j.aiaa.2022.09.007> [in English].
8. Markus Friedrich, Theresa Gerber, Jonas Dumler, Frank Döpfer. (2023). A system for automated tool wear monitoring and classification using computer vision. *Procedia CIRP*. Retrieved from <https://doi.org/10.1016/j.procir.2023.06.073> [in English].
9. Phat Thai, Sameer Alam, Nimrod Lilith, Binh T. Nguyen. (2022). An overview of Human Action Recognition in sports based on Computer Vision. *Journal Heliyon*. Retrieved from <https://doi.org/10.1016/j.heliyon.2022.e09633> [in English].
10. Nouriani, A. McGovern, R., Rajamani, R. (2023). Activity recognition using a combination of high gain observer and deep learning computer vision algorithms. *Intelligent Systems with Applications*. Retrieved from <https://doi.org/10.1016/j.iswa.2023.200213> [in English].
11. Masooma Memon. (2022). ANN vs CNN vs RNN: Neural Networks Guide. Retrieved from <https://levity.ai/blog/neural-networks-cnn-ann-rnn> [in English].
12. Zahra Elhamraoui. (2020). Introduction to convolutional neural network. Retrieved from <https://medium.com/analytics-vidhya/introduction-to-convolutional-neural-network-6942c189a723> [in English].
13. James, Le. (2018). The 5 Computer Vision Techniques That Will Change How You See The World. Retrieved from <https://heartbeat.comet.ml/the-5-computer-vision-techniques-that-will-change-how-you-see-the-world-1ee19334354b> [in English].
14. Jingfei Chang, Yang Lu, Ping Xue, Yiqun Xu, Zhen Wei. (2023). Iterative clustering pruning for convolutional neural networks. *Knowledge-Based Systems*. Retrieved from <https://doi.org/10.1016/j.knosys.2023.110386> [in English].
15. Vidushi Meel. ANN and CNN: Analyzing Differences and Similarities. Retrieved from <https://viso.ai/deep-learning/ann-and-cnn-analyzing-differences-and-similarities/> [in English].
16. Recurrent Neural Network Tutorial (RNN). (2022). Retrieved from <https://www.datacamp.com/tutorial/tutorial-for-recurrent-neural-network> [in English].
17. Kristina Host, Marina Ivašić-Kos. (2022). An overview of Human Action Recognition in sports based on Computer Vision. Retrieved from <https://doi.org/10.1016/j.heliyon.2022.e09633> [in English].
18. What is a recurrent network? Retrieved from <https://h2o.ai/wiki/recurrent-network/> [in English].