

УДК 004.72

DOI: 10.31673/2412-9070.2021.051115

Д. О. РОЗМАЇТИЙ, студент;

Н. В. ГРАБОВЕЦЬКИЙ, студент;

О. С. ЦВИК, студент;

А. В. ЛЕМЕШКО, доктор філософії,

Державний університет телекомунікацій, Київ

## ШТУЧНІ НЕЙРОННІ МЕРЕЖІ В СУЧАСНОМУ СВІТІ: ОСОБЛИВОСТІ ЗАСТОСУВАННЯ

*У статті розглянуто питання особливостей застосування штучних нейронних мереж у сучасному світі.*

*Штучні нейронні мережі широко використовуються в тих сферах, де традиційні комп'ютери є не надто ефективними. Наприклад, для розв'язання проблем, де необхідно, щоб система замість запрограмованих результатів вивчала, адаптувала та змінювала результати залежно від даних, які їй надходять. Нейронні мережі також знаходять істотне застосування, якщо йдеться про роботу з шумними або неповними даними.*

*Запропоновано шляхи використання штучних нейронних мереж у різних галузях промисловості.*

**Ключові слова:** нейронна мережа; комп'ютерний зір.

### Вступ

Для переважної більшості населення такі поняття, як глибоке навчання та штучний інтелект, все ще чужі. Більшість людей, які вперше стикаються з цими термінами, реагують зі змішаним почуттям скептицизму та остраху. В який саме спосіб можна змусити машини навчатися та виконувати роботу, призначену для людей? Що насправді пояснює всю галузь, яка прагне змусити машини поводитися так, як це роблять люди?

Найбільш проривним аспектом нейронних мереж є те, що після тренування вони навчаються самостійно. Отже, вони імітують людський мозок, який складається з нейронів, фундаментального будівельного матеріалу передавання інформації як людини, так і нейронної мережі.

Людські атрибути нейронних мереж і здатність виконувати завдання в нескінченних переставленнях та комбінаціях роблять їх унікальними для сучасних застосунків, що базуються на великих даних. Оскільки нейронні мережі також мають унікальну здатність (відому як нечітка логіка) розуміти неоднозначні, суперечливі або неповні дані, вони можуть використовувати контрольовані процеси за відсутності точних моделей.

### Основна частина

Галузь глибокого навчання набула популярності зі зростанням доступної оброблюваної потужності, місця для зберігання та великих обсягів даних. Замість використання традиційних моделей машинного навчання інженери штучного інтелекту переходять до моделей глибокого навчання. Там, де є багато даних, моделі глибокого навчання майже завжди перевершують традиційні моделі машинного навчання. Крім того, сфера глибокого навчання також швидко зростає. Багато дослідників докладають значних зусиль для розроб-

лення великих моделей глибокого навчання для розв'язання певних проблем. Незважаючи на те, що існують сотні альтернативних моделей глибокого навчання, їх можна згрупувати за основними категоріями.

Як доведено, деякі моделі є більш ефективними для реальних проблем [2]. Серед них нейронні мережі прямого поширення (*Feedforward Neural Network*), згорткові нейронні мережі (*Convolutional Neural Network*), рекурентні нейронні мережі (*Recurrent Neural Network*), трансформери (*Transformers*), автокодувальники (*Autoencoders*), генеративні змагальні мережі (*Generative Adversarial Network*).

*Нейронні мережі прямого поширення (FfNN) або багат шаровий перцептрон (MLP).* Нейронна мережа прямого поширення — це штучна нейронна мережа, в якій зв'язки між шарами не утворюють циклу [1, с. 17-20]. Це головна характеристика, яка робить нейромережі прямого поширення відмінними від рекурентних нейронних мереж. Нейромережа прямого поширення (або багат шаровий перцептрон) була першою і є найпростішим типом архітектури штучної нейронної мережі.

У цій мережі інформація рухається лише в прямому напрямку від вхідного шару, через приховані шари, до вихідного шару. У ній немає циклів.

Нейронні мережі прямого поширення, що використовуються сьогодні, є вдосконаленими версіями оригінальних та простих багат шарових перцептронів (MLP). Ці терміни часто використовують як взаємозамінні. У сучасних нейромережах постійного поширення часто застосовують складніші функції активації, ніж ступенева (0 або 1), такі як випрямляч (ReLU), сигмоїда, гіперболічна функція та нормована експоненційна функція. Сучасні глибокі нейронні мережі зазвичай

використовують один із методів градієнтного спуску для оптимізації.

Нейронні мережі прямого поширення — відмінна архітектура для початку вивчення моделей глибокого навчання. Попри те, що вони мають свої недоліки, нейромережі прямого поширення можуть бути використані для побудови програм розпізнавання мови, зображень та машинного перекладу. Вони надзвичайно ефективні в аналітиці даних.

**Згорткові нейронні мережі (CNN).** Однією з найпотужніших моделей глибокого навчання під наглядом є згорткові нейронні мережі [1, с. 40-42]. Остаточна структура CNN насправді дуже схожа на нейронні мережі прямого поширення, де є нейрони з вагами та упередженнями. Крім того, як і в нейронних мережах прямого поширення, у згорткових нейронних мережах є функція втрат (наприклад, перехресна ентропія або softmax) та оптимізатор (наприклад, оптимізатор Адама). Крім того, у CNN існують також згорткові шари, агрегувальні шари та повноз'єднані шари. Згорткові нейронні мережі зазвичай використовують для завдань комп'ютерного зору, зокрема класифікації зображень, хоча є й інші сфери застосування, такі як оброблення природної мови та прогнозування часових рядів.

**Згортковий шар** — це перший шар, на який витягуються функції із зображень наборів даних. Оскільки пікселі пов'язані лише з сусідніми та ближніми пікселями, згортка дає можливість зберегти взаємозв'язок між різними частинами зображення. Конволюція (згортка) — це переважно фільтрація зображення за допомогою меншого піксельного фільтра для зменшення розміру зображення без втрати взаємозв'язку між пікселями. Під час застосування згортки до зображення  $5 \times 5$ , використовуючи фільтр  $3 \times 3$  з кроком  $1 \times 1$  (зсув одного пікселя на кожному кроці), отримуємо зображення  $3 \times 3$  (зменшення складності на 64%).

У процесі побудови згорткової нейронної мережі після кожного згорткового шару зазвичай вставляють **агрегувальні шари**, щоб зменшити просторовий розмір подання для скорочення кількості параметрів, що послаблює обчислювальну складність. Крім того, агрегувальні шари також допомагають розв'язати проблему переобладнання. Здебільшого розмір об'єднання вибирають для зменшення кількості параметрів, обравши максимум, середнє значення або значення суми всередині цих пікселів. **Повноз'єднані шари.** Повністю підімкненою мережею є нейронна мережа прямого поширення, де кожен параметр зв'язаний між собою, щоб визначити справжнє відношення та вплив кожного параметра на позначки.

**Рекурентні нейронні мережі.** Рекурентні нейронні мережі (RNN) — це клас штучних нейрон-

них мереж, де зв'язки між нейронами утворюють часову послідовність [1, с. 271-274]. RNN здатні фіксувати динамічну часову інформацію (тимчасова пам'ять). Рекурентні нейронні мережі походять від нейронних мереж прямого поширення, але вони пропонують значно більше. RNN використовують попередню інформацію, відкладаючи її в пам'яті, яка зберігається як «стан» у нейроні рекурентних нейронних мереж.

Існує багато різних варіантів рекурентних нейронних мереж, але є такі типи RNN, якими користуються частіше:

- ванільні (прості) RNN;
- мережі довгої короткочасної пам'яті (LSTM) [1, с. 292-295];
- мережі з вентиляним рекурентним вузлом (GRU) [1, с. 295-297].

**Трансформери.** Трансформери — це моделі глибокого навчання, які було введено 2017 року. Трансформери переважно використовуються для завдань з оброблення природних мов (NLP).

Трансформери було розроблено як відповідь на обчислювальну складність і повільність рекурентних нейронних мереж. Трансформери можуть обробляти послідовні завдання даних, зокрема машинний переклад, передбачення часових рядів або узагальнення тексту.

Основна перевага трансформерів полягає в тому, що, на відміну від RNN, вони не потребують послідовного оброблення даних. Якщо вхідними даними є показники продажів у часових рядах, трансформеру не потрібно обробляти ранні дати перед пізніми. Отже, трансформери дають змогу набагато більше розпаралелювати, ніж RNN, і, отже, скорочують час навчання.

Оскільки вони значно ефективніші щодо швидкості навчання, трансформери стали вибором дослідників NLP і замінюють альтернативи RNN. Наприклад, популярні попередньо навчені моделі NLP, такі як BERT і GPT-3, є моделями трансформерів, які можна використовувати для завдань NLP з передавальним навчанням.

**Автокодувальники.** Автокодувальники — це архітектури нейронних мереж, які складаються з двох підмереж: шифратора та дешифратора, прив'язаних один до одного прихованим простором [1, с. 70-73]. Уперше автокодувальники було розроблено Джеффрі Хінтоном, одним із найбільш шанованих учених у спільноті штучного інтелекту, та групою PDP у 1980-х. Хінтон та група PDP мали на меті розв'язати проблему «зворотного поширення помилки без учителя», тобто нагляду без учителя, використовуючи дані як викладач.

Автокодувальники складаються з мережі кодувальника, яка приймає дані ознак та кодує їх, щоб вони помістилися у прихований простір та мережу дешифратора. Дешифратор використовує ці зако-

довані дані (тобто код) і перетворює їх знову в дані ознак. У шифраторі модель навчається ефективно кодувати дані, щоб дешифратор міг перетворити їх назад в оригінал. Отже, важливою частиною навчання автокодувальника є створення оптимізованого прихованого простору.

Здебільшого кількість нейронів у прихованому просторі набагато менша, ніж у вхідному та вихідному шарах, але це не має так бути. Існують різні типи автокодувальників, зокрема недоповнені, переповнені, розріджені, знешумлювальні, скорочувальні та варіаційні автокодувальники.

У процесі побудови автокодувальника стандартною практикою є проектування шифратора та створення оберненої версії цієї мережі як шифратора цього автокодувальника. Отже, поки існує зворотна залежність між шифратором та мережею шифратора, можна вільно додавати будь-який шар до цих підмереж. Наприклад, під час роботи з даними зображень, безсумнівно, знадобляться згорткові та агрегувальні шари. Водночас, якщо це робота з даними послідовності, то стануть у пригоді одиниці LSTM, GRU або RNN. Важливим моментом тут є те, що з'являється можливість будувати будь-що.

**Генеративні змагальні мережі.** Генеративні змагальні мережі пропонують революційні генеративні вирішення, які можуть бути використані в різних сферах, зокрема моделювання білків. Є два агенти, які конкурують між собою у звичайній структурі GAN: генератор та дискримінатор [1, с. 438-449]. Вони можуть бути спроектовані з використанням різних мереж (наприклад, згорткових нейронних мереж (CNN), рекурентних нейронних мереж (RNN) або нейромереж прямого поширення (FfNN)).

Як приклад, для завдання генерації зображень генератор просить згенерувати зображення, не надаючи йому жодних додаткових даних. Одночасно збирають реальні зображення з набору даних на дискримінатор і просять його вирішити, чи є зображення, сформовані генератором, справжніми чи ні. Спочатку генератор видає шуми, які дискримінатор негайно позначить як підроблені. Зрештою генератор навчиться обманювати дискримінатор, отримавши достатній відгук від дискримінатора. У результаті дістаємо генеративну модель, яка може дати дуже реалістичні результати.

**Застосування штучних нейронних мереж у сучасному світі.** Глибоке навчання — це наступна еволюція машинного навчання. Машинне навчання та глибоке навчання є одночасно підмножиною штучного інтелекту. Глибоке навчання — це розвинена та вдосконала фаза машинного навчання. Глибоке навчання вчиться і розуміє речі за допомогою штучної нейронної мережі, яка

дуже схожа на людський мозок і дає можливість машині аналізувати дані в структурі так само, як це роблять люди. Основна перевага глибокого навчання полягає в тому, що воно не вимагає від програміста інструкцій, що саме слід робити з даними. Це стає можливим завдяки надзвичайній кількості даних, що збираються та споживаються.

Суть глибокого навчання зводиться до того, що в нього є досить швидкі комп'ютери та достатньо даних, аби тренувати великі нейронні мережі. Глибоке навчання стало популярним завдяки навчанню під контролем. Глибоке навчання чудово працює на проблемних доменах, де вхідні дані та здебільшого навіть вихідні дані є аналоговими. А отже, це не кілька величин у табличному форматі, а натомість це зображення піксельних даних, документи текстових даних або файли аудіоданих. Це трубопровід модулів, усі з яких можна глибоко навчити, оскільки він має кілька етапів у процесі розпізнавання об'єкта, і всі ці етапи є частиною навчання.

**Комп'ютерний зір.** Сучасні камери, якими б досконалими вони не були, не можуть визначити предмет, що бачать на об'єктиві. Людський мозок використовує минулий досвід як контекст для класифікації та визначення того, що бачить людське око. Машини не мають такої можливості [3].

Дисципліна, яка дає змогу принаймні частково відтворити систему використання людського мозку знань для осмислення предметів, називається глибоким навчанням. У конкретному випадку з комп'ютерним зором переважно використовують згорткові нейронні мережі (CNN).

CNN роблять це, розбиваючи кожне зображення на невеликі матриці пікселів та слайд-фільтри (математичні об'єкти), щоб виконувати операції згортки для виявлення закономірностей. На першому рівні, подібно до людського мозку, CNN визначають такі речі, як грубі криві та межі зображення. Через кілька згорток вони починають складати між собою поверхні, інформацію про глибину, шари, розриви у візуальних просторах і, нарешті, починають розбирати такі предмети, як обличчя, одяг, риби, машини, тварини тощо.

На початку машина здебільшого виходить із ладу, оскільки всі значення її фільтрів рандомізовані. Проте вона продовжує порівнювати свої випадкові результати з фактичними з позначеного набору даних та використовувати функції виправлення помилок та втрат, поступово збільшуючи точність.

CNN виявилися чудовими для визначення особливостей нерухомих зображень, але вони повністю втрачають результативність в разі оброблення серії кадрів зображень, тобто відео. Вони не здатні ідентифікувати елементи, які можуть змінитися з часом, і зрозуміти контекст прогресування



зображень, що є важливим фактором для правильного маркування відео.

Отже, для оброблення відео фахівці з комп'ютерного зору спираються на роботу CNN, а потім уводять у рівняння інший тип алгоритмів — рекурентні нейронні мережі (RNN). Вони подають вихідні дані із згорткової мережі в RNN, оскільки останні оснащені часовою структурою кожного відеокліпу.

Ключова відмінність CNN від RNN полягає в тому, що перші мають справу з кожною матрицею пікселів незалежно, тоді як другі здатні «запам'ятовувати» вже оброблені ними дані та приймати рішення на основі нагромаджених знань.

Тут також варто згадати, що нейронні мережі, які часто можуть бути надмірно пристосованими, можуть отримати вигоду з передавального навчання, тобто використання моделей, які були навчені для інших цілей, як вихідний пункт для моделей, які працюють над новим завданням. Наприклад, алгоритм, який ефективно виявляє тварин, можна навчити далі розрізняти людей тощо. Цей метод може допомогти набагато легше отримати навчальний набір даних достатнього розміру.

Комп'ютерний зір швидко застосовується для різних цілей у галузі охорони здоров'я, сільського господарства, автомобільної промисловості тощо [4].

Медична візуалізація зростає роками, і численні стартапи в галузі охорони здоров'я співпрацюють із відомими постачальниками обладнання для створення найсучасніших засобів комп'ютерного зору. Донедавна одним із найпопулярніших випадків використання було застосування CNN для виявлення захворювань за допомогою МРТ. Але зараз компанії, такі як Arterys, отримали дозвіл від FDA застосовувати глибоке навчання в клінічних умовах.

Техніка безпілотних літальних апаратів також бурхливо розвивається, і в результаті прогресу в галузі витрати на придбання величезних наборів аерофотознімків значно зменшилися порівняно з кількома минулими роками. Це, у поєднанні з недавніми досягненнями в галузі машинного навчання, надає багато перспектив для аграрного бізнесу. Комп'ютерний зір може допомогти фермерам виявляти хвороби сільськогосподарських культур, прогнозувати врожайність і, загалом, автоматизувати трудомісткі процеси ручного польового обстеження.

Окрім автомобілів із ручним керуванням, в автомобільній промисловості широко застосовується використання комп'ютерного зору. Мережі Tesla за допомогою камери аналізують вихідні зображення для семантичної сегментації, виявлення об'єктів та оцінювання глибини монокуляра.

Мережі з висоти пташиного польоту беруть відео з усіх камер для виведення плану дороги, статичної інфраструктури та 3D-об'єктів безпосередньо у вигляді зверху вниз. Мережі вчаться на найскладніших та різноманітніших сценаріях у світі, дістаючи інформацію з майже 1 млн автомобілів у режимі реального часу. Повна побудова нейромереж автопілоту охоплює 48 мереж, на підготовку яких потрібно більш як 70 000 год тренування графічного процесора. Разом вони видають 1000 різних тензорів (прогнозів) на кожному кроці часу.

**Автоматизовані переклади.** Автоматизовані переклади існували ще до того, коли почали використовувати глибоке навчання. Але глибоке навчання допомагає машинам робити розширені переклади з гарантованою точністю, якої бракувало в минулому. Крім того, глибоке навчання також допомагає в перекладі, отриманому із зображень — чогось абсолютно нового, що не могло б бути можливим за допомогою традиційної текстової інтерпретації.

**Формування тексту.** Зараз машини мають можливість генерувати новий текст із нуля. Вони можуть навчитися пунктуації, граматиці та стилю фрагмента тексту та писати ефективні новини. Робожурналісти, що базуються на моделях глибокого навчання, готують точні звіти про спортивні змагання вже щонайменше протягом трьох років. І вміння не обмежується виключно написанням звітів. Генерація тексту на основі штучного інтелекту повністю підготовлена для вирішення складності думок. На цей час генерація тексту допомагає створювати записи майже про все — від дитячих віршів до наукових тем.

**Перетворення голосу в текст.** Сьогодні однією з найпоширеніших функцій смартфонів є перетворення голосу в текст. Просто натискання кнопки або вимовлення певної фрази (наприклад, «Ok Google») дає можливість розпочати розмову з телефоном, і телефон перетворює звук у текст.

**Розумні особисті помічники.** Коли технологія передавання голосу до тексту стає настільки точною, що на неї можна покладати основні розмови, вона перетворюється на інтерфейс керування для нового покоління персональних помічників. Серед них Siri, Google Assistant (Google Now), Amazon Alexa, Microsoft Cortana, Bixby, Voice Mate, Аліса. Вони можуть виконувати пошук в інтернеті, встановлювати нагадування, інтегруватися у ваш календар тощо.

### Висновки

У статті було розглянуто особливості застосування штучних нейронних мереж у сучасному світі. Зокрема, досліджено основні моделі штучних нейронних мереж, що активно застосовуються, та сфери їх використання.

Утім це лише кілька поверхневих варіантів застосування штучних нейронних мереж. Уже зараз можна натренувати моделі для досягнення точності, яка часом конкуруватиме з людськими можливостями розпізнавання зображень, а численні компанії, великі та малі, знайдуть способи використовувати ці досягнення для зменшення операційних витрат та впорядкування бізнес-процесів. Подальше зростання моделей глибокого навчання надасть набагато більше методів використання штучного інтелекту в сучасному світі.

#### Список використаної літератури

1. **Aggarwal C.** *Neural Networks and Deep Learning*. Yorktown Heights, USA: Springer International Publishing AG, 2018. 512 p.

2. **Top 6 Deep Learning Models You Should Master for Killer AI Applications** [Електронний ресурс]. URL:

<https://towardsdatascience.com/top-6-deep-learning-models-you-should-master-for-killer-ai-applications-13c7dfa68a3>.

3. **Computer Vision Applications: How Real-Time Image Processing is Reshaping Industries and How Your Business Can Leverage It** [Електронний ресурс]. URL:

<https://perfectial.com/blog/computer-vision-applications/>.

4. **Neural Networks: Applications in the Real World** [Електронний ресурс]. URL:

<https://www.upgrad.com/blog/neural-networks-applications-in-the-real-world/>.

Д. О. Розмайтій, Н. В. Грабовецкий, А. С. Цвык, А. В. Лемешко

#### ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ В СОВРЕМЕННОМ МИРЕ: ОСОБЕННОСТИ ПРИМЕНЕНИЯ

В статье рассматривается вопрос об особенностях применения искусственных нейронных сетей в современном мире.

Искусственные нейронные сети находят широкое применение в областях, где традиционные компьютеры не слишком эффективны. Например, для решения проблем, где необходимо, чтобы система вместо запрограммированных результатов изучала, адаптировала и изменяла результаты в зависимости от поступающих ей данных. Нейронные сети также находят серьезное применение, когда говорят о работе с шумными или неполными данными.

Предложены пути применения искусственных нейронных сетей в различных отраслях промышленности.

**Ключевые слова:** нейронная сеть; компьютерное зрение.

D. O. Rozmaityi, N. V. Hrabovetskyi, O. S. Tsvyk, A. V. Lemeshko

#### ARTIFICIAL NEURAL NETWORKS IN THE MODERN WORLD: FEATURES OF APPLICATION

This article discusses the peculiarities of the use of artificial neural networks in the modern world.

Artificial neural networks are widely used in areas where traditional computers do not work very well. For example, to solve problems where instead of programmed results it is necessary that the system studies, adapts and changes the results depending on the data it receives. Neural networks are also widely used when talking about working with noisy or incomplete data.

Due to the ability to learn and adapt, similar to the brain, neural networks form a complete basis and are used in artificial intelligence and, consequently, in machine learning algorithms.

As we collect more and more data every year, it makes sense to use deep learning models. In addition, the field of deep learning is also developing rapidly. Many researchers are working hard to develop excellent models of deep learning for specific problems.

The use of deep neural networks and other machine learning technologies to solve long-standing problems is evolving rapidly and promises to shape the future of technology. Modern technologies use both controlled and uncontrolled methods. They are likely to become central to practice, and will soon become as widespread and inconspicuous as other technologies we have integrated into everyday use. The scale of the problems that machine learning can help solve is enormous and is likely to evolve rapidly. Integrating machine learning into the daily workflow can expand our capabilities and make modern problem-solving methods more efficient, more focused on higher-order tasks.

Neural networks solve problems that require pattern recognition. They are good for pattern recognition, classification and optimization. These include handwriting recognition, face recognition, speech recognition, text translation, credit card fraud detection, medical diagnostics and huge data solutions. You can use it to find connections between templates, convert one type of data to another, and create associations or generalizations between different entities.

The article will consider the application of artificial neural networks in various industries.

**Keywords:** neural network; computer vision.