

DOI: <https://doi.org/10.36910/6775-2524-0560-2021-43-36>

УДК 004.45, 004.08

Тулашвілі Юрій Йосипович, д.пед.н., професор,

<https://orcid.org/0000-0002-0780-9529>

Лук'янчук Юрій Анатолійович, к.т.н., ст.викладач

<https://orcid.org/0000-0001-9690-6197>

Луцький національний технічний університет

## ВИКОРИСТАННЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ НА ОСНОВІ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ОБРОБКИ ЗОБРАЖЕНЬ

Тулашвілі Ю.Й., Лук'янчук Ю.А. Використання програмного забезпечення на основі штучного інтелекту для обробки зображень. В даній статті описуються проблеми обробки графічної інформації, шляхи їх вирішення та перспективи розвитку програмного забезпечення на основі штучного інтелекту. Також описано раціональне використання різного типу нейромереж в повсякденному житті, їх якісний вплив на розвиток підприємств різних галузей науки та виробництва та інші важливі шляхи застосування програмних продуктів на базі ШІ.

**Ключові слова:** програмне забезпечення, нейромережі, обробка графічної інформації.

Tulashvili Y., Lukianchuk I. Use of artificial intelligence software for image processing. This article describes the problems of graphic information processing, ways to solve them and prospects for the development of software based on artificial intelligence. It also describes the rational use of different types of neural networks in everyday life, their qualitative impact on the development of enterprises in various fields of science and production and other important ways to use software products based on AI.

**Keywords:** software, neural networks, graphic information processing.

Тулашвили Ю.И., Лукьянчук Ю.А. Использование программного обеспечения на основе искусственного интеллекта для обработки изображений. В данной статье описываются проблемы обработки графической информации, пути их решения и перспективы развития программного обеспечения на основе искусственного интеллекта. Также описано рациональное использование различного типа нейронных сетей в повседневной жизни, их качественное влияние на развитие предприятий различных отраслей науки и производства и другие важные пути применения программных продуктов на базе ИИ.

**Ключевые слова:** программное обеспечение, нейросети, обработка графической информации.

Сьогодні компанії будь-якого профілю мають необхідність обробки великої кількості даних та зображень. Від якості і швидкості їх аналізу залежить ефективність прийняття рішень і рівень підтримки клієнтів. Впоратися зі зростанням інформаційних потоків допомагають технології штучного інтелекту. В їх основі лежать алгоритми глибокого машинного навчання, що поєднують в собі нейромережі певних типів. Однак в сучасному світі існує проблема обробки отриманої графічної інформації, оскільки вона потребує додаткового редагування. Тому використання програмного забезпечення на основі штучного інтелекту, що покращує роздільну здатність зображень, потребує додаткового вивчення.

Згідно із дослідженнями [1,2,3] вважається, що штучні нейронні мережі ANN - це проривна технологія, однак на рівні алгоритмів вони були описані ще в другій половині минулого століття. Наразі з'явилася технічна можливість реалізувати їх в кінцевих продуктах - людство накопичило достатні обсяги інформації і створило засоби для її швидкого аналізу. До теперішнього часу найбільшого поширення набули такі види нейромереж: - нейронні мережі CNN, що імітують роботу зорової кори головного мозку і частково виконують функцію абстрактного мислення. Вони прекрасно справляються із завданням розпізнавання зображень, а їх обчислення можна розподілити на графічних процесорах, що дозволяє створювати відносно дешеві апаратні платформи з елементами ШІ.

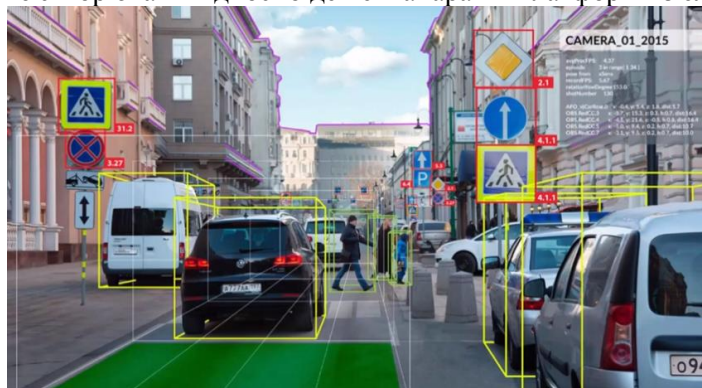


Рис. 1. Розпізнавання об'єктів за допомогою CNN

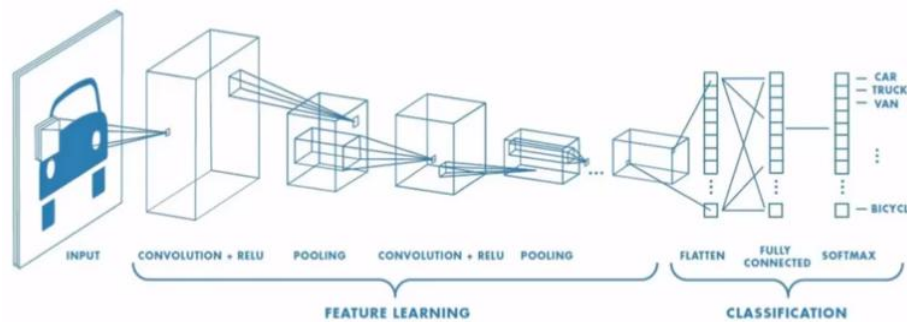


Рис. 2. Класифікація об'єктів за допомогою CNN ([https://youtu.be/EQS7zjyaV\\_k](https://youtu.be/EQS7zjyaV_k))

CNN застосовуються в системах машинного зору безпілотних автомобілів, комерційних дронів, роботів, а також в охоронному відеоспостереженні. Людина щодня використовує CNN, якщо ввімкнено в налаштуваннях смартфона розблокування за допомогою розпізнавання особи.

- нейронні мережі RNN, мають короткочасну пам'ять, за рахунок чого легко аналізують послідовності довільної довжини. RNN розбивають потік даних на елементарні частини і оцінюють взаємозв'язки між ними. Ці алгоритми знайшли основне застосування в розпізнаванні рукописного тексту й мови. Наприклад, під час пошуку мелодії на слух в Shazam, розмови з Siri, Google Now або Алісою від «Яндекса», роблячи замітки від руки для Cortana - на хмарних платформах працюють нейромережі.

- мережі з довготривалою і короткочасною пам'яттю LSTM, що стали наступним етапом розвитку RNN. За допомогою них відбувається прогнозування змін будь-якої величини (наприклад, біржових курсів або купівельного попиту) шляхом екстраполяції. Також їх застосовують для глибокого аналізу природної мови. Наприклад, Google використовує LSTM в персональному помічнику і системі машинного перекладу Google Translate.

- керовані блоки GRU - нова модифікація RNN, що з'явилася в 2014 році. Їх використовують для синтезу мови, яка володіє емоційним забарвленням і звучить як справжня. Наприклад, в тестах сервісів Google Duplex і Microsoft Xiaoice люди не змогли відрізнити розмову ботів від живих співрозмовників. Це призвело до того, що Xiaoice дозволила Microsoft зміцнитися на азіатському ринку, де розвиток компанії завжди стримувався мовним бар'єром.

- глибокі нейронні мережі DNN - будь-яка мережа більш ніж з трьома шарами. Вони лежать в основі механізмів глибокого машинного навчання, знаходячи неявні взаємозв'язки між різнорідними даними. Наприклад, пошук кореляцій між розвитком захворювань і різними потенційними факторами у величезних масивах наукових статей за допомогою IBM Watson.

- генеративні мережі GAN - це комбінація нейромереж, одна з яких генерує варіанти, а інша відсіває їх. Таке поєднання дозволяє реалізувати машинне навчання без вчителя, що підвищує його автономність. Наприклад, PixelDTGAN генерує окремі зображення одягу, взуття та аксесуарів для каталогів онлайн-магазинів. В якості вхідних даних використовуються фотографії, на яких ці предмети гардеробу демонструють фотомоделі. Поки зйомку для каталогів одягу вважається досить витратною частиною електронної комерції, але цілком можливо, що в найближчі роки нейромережі дозволять швидко проілюструвати каталог, навіть не залучаючи фотографа. Обробка фотографій теж забирає багато часу. На практиці це відбувається набагато швидше, додаючи і прибираючи графічні об'єкти за допомогою іншої нейромережі - IBM GANPaint. Подібна їй нейромережа DRAGAN вже застосовується для автоматичного відтворення персонажів аніме і мультфільмів. Вона дозволяє пришвидшити вихід нових серій і утримати аудиторію розважальних каналів, не перевантажуючи аніматорів колосальним об'ємом роботи. Також, GAN дозволяє анімувати тривимірну модель людини, переносячи на неї рух іншої людини в реальному часі. Перші результати виглядають не дуже переконливо, однак особливність нейромережі в тому, що вона покращується з кожною новою порцією даних.

Google використовує нейронні мережі в усіх своїх сервісах - від онлайн-пошуку і фільтрації пошти до підбору роликів на YouTube і перекладу текстів. На відміну від старих моделей машинного перекладу, нейромережі Google оперують практично необмеженими об'ємами даних для розпізнавання мовних шаблонів. Вони можуть виявляти більше семантичних зв'язків і покращувати результат з кожним разом за допомогою постійного навчання.

З 2001 року в Google став можливий пошук по зображеннях, проте спочатку для цього застосовувався простий аналіз метаданих. Тобто фактично пошук відбувався по текстовому опису картинки, який часто був неточним. Сучасна версія цього сервісу використовує систему автоматичного

опису графічних файлів. Вона побудована на комбінації нейромереж, тому в ній можна шукати подібні зображення, сортувати їх по типам і навіть знаходити графічні файли без текстової анотації.

Google Cloud Speech-to-Text став необхідною інновацією для великих компаній, що працюють з клієнтами з різних країн. Вихід за межі регіону завжди потребував високих вимог до організації технічної підтримки. Якщо раніше для цього доводилося відкривати регіональні філії або наймати перекладачів, то тепер в рішення для колл-центрів може бути інтегрований цей сервіс. Він підтримує інтерфейс прикладного програмування і допомагає перетворювати запис розмов в текст, застосовуючи моделі нейронних мереж. В даний час він розпізнає понад 120 мов і продовжує освоювати нові.

LinkedIn використовує нейронні мережі спільно з лінійними текстовими класифікаторами для аналізу всіх записів, які розповсюджуються через свою платформу. Таке поєднання допомагає ефективніше виявляти спам і контент образливого змісту, а також підбирати кращі персональні рекомендації.

Facebook застосовує подібний алгоритм аналізу тексту. Він називається DeepText і переважно використовується для таргетування рекламних оголошень. Додатково Facebook використовує нейронні мережі для розпізнавання осіб і обробки зображень.

Tesla використовує набір DNN у фірмовій системі автопілота. Вони розпізнають всі об'єкти навколо машини в реальному часі, класифікують їх і визначають характеристики. Дев'ята версія автопілота стала в чотири рази складніше восьмої за рахунок збільшення потоку даних. Одних камер стало вісім штук, причому їх роздільна здатність теж збільшилася [4].



Рис. 3. Розпізнавання об'єктів за допомогою DNN в системі автопілота від Tesla

Антивірусні розробники застосовують різні нейронні мережі для пошуку нових загроз в галузі інформаційної безпеки. Вони допомагають розпізнавати невідомі модифікації шкідливих файлів і нові види мережевих атак. Із презентацією системи ICSP Neural від Symantec можна ознайомитись за посиланням [5].

Зовсім недавно штучні нейронні мережі були доступні тільки великим компаніям, але за останні роки з'явилося безліч доступних рішень, готових засобів розробки і хмарних сервісів, які легко адаптувати до своїх завдань. Наприклад, IBM випустила бета-версію Neural Network Modeler для Watson Studio. Цей інструмент дозволяє використовувати візуальний підхід при розробці архітектур DNN для обробки зображень, тексту і аудіофайлів, не маючи навичок програмування.

У березні 2018 року Microsoft представила Project Brainwave - програмно-апаратний комплекс на програмованих матрицях (FPGA) виробництва Intel, який полегшує нейромережам завдання розпізнавання і класифікації об'єктів в реальному часі. Зараз ця платформа використовується в пошуковикі Bing, прискорюючи обробку запитів майже на порядок. Також вона тестується одним з партнерів Microsoft в автоматизованій системі оптичної дефектоскопії для прискорення конвеєрного складання.

У листопаді 2018 року Intel представила Neural Compute Stick 2 - нове покоління компактних прискорювачів для нейромереж. Володіючи розмірами флешки, Intel NCS2 підключається до USB-порту будь-якого пристрою. Для складних завдань можна задіяти одразу кілька таких прискорювачів, а їх число - змінити в будь-який момент, що спрощує поетапне впровадження технологій ШІ.

На практиці, досить часто виникає необхідність обробки отриманих зображень, тому розробка та використання програмного забезпечення на основі штучного інтелекту є актуальним та доцільним завданням, що потребує глибокого аналізу.



В дизайнерів із різних галузей виникає проблема, коли необхідне зображення із інтернету має низьку роздільну здатність і погану якість, а замовник продукту потребує саме цю картинку. Тому покращення відбувається за допомогою засобів ретушування, наприклад Photoshop, однак немає можливості детального редагування або додавання пікселів. Та й не всі фахівці володіють відповідними навичками. В цьому випадку доцільно використовувати програмне забезпечення на основі штучного інтелекту.

Нейромережі вже давно і цілком успішно справляються із ретушуванням зображень. Також існує велика кількість програм алгоритмічного редагування. Проте лідером у цій сфері є пакет програмного забезпечення від Topaz Labs. Для збільшення роздільної здатності фотографій є програма під назвою Gigapixel AI. Вона не тільки збільшує роздільну здатність, а й реально покращує якість самого зображення і тому є необхідністю вивчення даного продукту.



Рис. 4. Обробка зображення пейзажу за допомогою Topaz Gigapixel AI

В основі алгоритмів програми лежить цілий комплекс нейромереж. Програма розбирає фото на пікселі і сортує їх за кольором, потім формується маска в потрібній роздільній здатності і далі система схематично додає пікселі потрібних кольорів на наявне зображення, таким чином буквально «домальовуючи» його до необхідного покращення [6, 7].

Topaz Gigapixel AI буде корисна для дизайнерів у різних сферах, де використовують роботу із зображеннями. Наприклад, якщо:

- кадр зроблений на старий фотоапарат, телефон або фотоплівку;
- сталася невелика помилка з фокусом;
- фото було дуже урізано, через що постраждала деталізація;
- знімок потрібно роздрукувати великим форматом (більше 1м2).

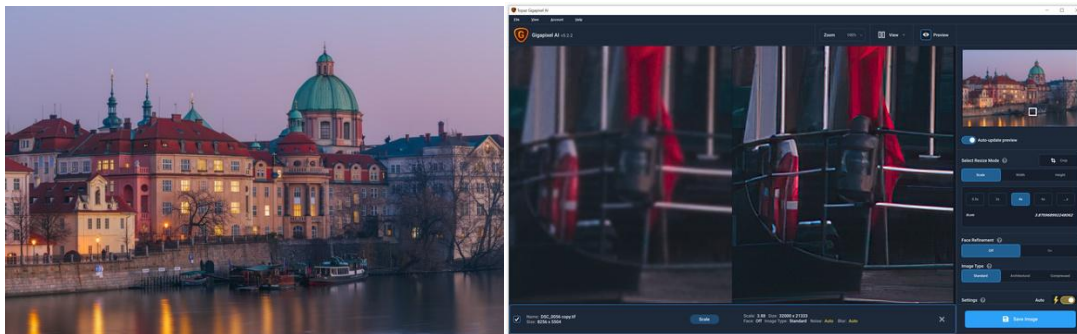


Рис. 5. Обробка зображення корми човна за допомогою Topaz Gigapixel AI

Topaz Gigapixel AI не просто «розтягує» зображення (як це відбувається в Adobe Photoshop), а розпізнає деталі і текстури, роблячи їх більш чіткими. Так знімок можна збільшити, не втративши в деталізації. Можна піти зворотним шляхом: збільшивши недостатньо якісний знімок, покращити деталізацію.

Однак завантажувати відверто нерізкі, зіпсовані знімки сенсу немає - на них не вийде розпізнати достатню кількість деталей, результат буде посереднім.

По прев'ю, що видане програмою, по картинці до/після видно, наскільки ефективно працює програма.

Програма обробляє зображення та показує прев'ю. Можна вибрати кратність збільшення знімка. Краще ставити x2 або x4, цього достатньо, щоб виявити дрібні деталі і фактури на фото. Чим більше кратність, тим довше програма буде опрацьовувати результат. Крім кратності збільшення, можна

більше нічого і не вибирати, у Toraz GigaPixel AI дуже лаконічний інтерфейс, програма працює практично повністю автоматично.

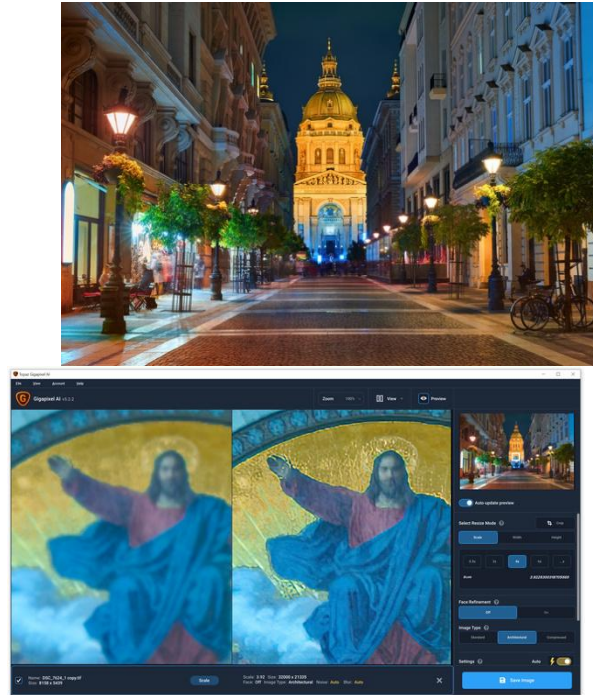


Рис. 6. Обробка зображення будівлі за допомогою Toraz GigaPixel AI

Рекомендується проводити подальшу обробку зображень за допомогою інструментів Photoshop, щоб провести ресайз та зменшити розмір файлу зображення приблизно в 2-3 рази, не втративши в деталізації. Також цей інструмент дозволяє сховати дефекти роботи Toraz GigaPixel AI (вони іноді з'являються, якщо збільшувати фотографії в 4 або 6 разів).

На сьогоднішній день створено багато нових інструментів для редагування зображень. При ручній обробці в Photoshop на картинку витрачається багато часу. А це неприпустимо в сучасному світі: поки один дизайнер робить один кадр, конкуренти зроблять сто. Тому необхідно вивчати та розробляти програмне забезпечення на основі штучного інтелекту, що пришвидшує обробку зображень. Особливість програм на основі ШІ в тому, що не завжди контролюється результат від початку та до кінця, однак в 95% він виходить відмінним. А для решти 5% випадків важливо вміти працювати з класичними інструментами Adobe Photoshop. Така синергія класичних методів обробки та сучасних технологій дозволяють робити обробку не тільки швидко, але й якісно. Нейронні мережі можуть допомогти компаніям розвиватись відразу по декількох напрямках: оптимізувати бізнес-логістику, покращити прогнозування та взаємодію з клієнтами. В кінцевому рахунку вони дають можливість підвищити конкурентоспроможність за рахунок більш глибокого аналізу всіх доступних даних. Очікування клієнтів постійно зростають, і вони вибирають сервіси на базі ШІ, оскільки вони швидше і повніше задовольняють їхні запити.

#### Список бібліографічного опису

1. Искусственные нейронные сети и приложения: учеб. пособие / Ф.М. Гафаров, А.Ф. Галимянов. – Казань: Изд-во Казан. ун-та, 2018. – 121 с.
2. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2002. – 1072 с.
3. <https://dx.media/articles/how-it-works/neyroseti-v-biznese-zachem-i-dlya-chego/>
4. <https://youtu.be/7ztK5AhShqU>
5. <https://youtu.be/rung43Zsbpw>
6. <https://root-nation.com/ru/soft-all/howto-ru/ru-kak-uvlichit-razreshenie-foto-gigapixel-ai/>
7. <https://prophotos.ru/lessons/22466-neyronnye-seti-na-sluzhbe-fotografii/3>

#### References

1. Artificial neural networks and applications: textbook. manual / F.M. Gafarov, A.F. Galimyanov. - Kazan: Kazan Publishing House. University, 2018. -- 121 p.
2. Gonzalez R., Woods R. Digital image processing. - M.: Technosphere, 2002. -- 1072 p.
3. <https://dx.media/articles/how-it-works/neyroseti-v-biznese-zachem-i-dlya-chego/>
4. <https://youtu.be/7ztK5AhShqU>
5. <https://youtu.be/rung43Zsbpw>
6. <https://root-nation.com/ru/soft-all/howto-ru/ru-kak-uvlichit-razreshenie-foto-gigapixel-ai/>
7. <https://prophotos.ru/lessons/22466-neyronnye-seti-na-sluzhbe-fotografii/3>