

Н.О. Кушнір, асистент
Т.М. Локтікова, ст. викладач
А.В. Морозов, к.т.н., доц.
В.О. Юрченко, магістрант

Державний університет «Житомирська політехніка»

Використання згорткових нейронних мереж у задачах розпізнавання та класифікації об'єктів зображень

Досліджується застосування штучних нейронних мереж у задачах розпізнавання та класифікації об'єктів зображень. Розглядається класична задача класифікації об'єктів на зображенні, а саме визначення статі особи за її обличчям. Це зумовлено тим, що існує вдалий набір даних, який складається з 47009 зображень обличчів чоловіків та жінок для навчання та 11649 зображень обличчів для тестування штучної нейронної мережі. Запропоновано використання згорткової нейронної мережі. Такий підхід дозволяє зменшити обсяг інформації, що зберігається у пам'яті, а також ієрархічно виділити та агрегувати ознаки вхідних даних. Згорткова нейронна мережа складається з декількох блоків згорткових й агрегувального шарів, шару вирівнювання, шарів повнозв'язних нейронів, вихідного нейрона. Пороговою активаційною функцією для всіх нейронів, окрім вихідного, обрано функцію «ReLU». Активаційною функцією вихідного нейрона є сигмоїдальна. Здійснено побудову нейронної мережі, її навчання та тестування із застосуванням бібліотеки «TensorFlow», API «Keras.NET», а також розробленої бібліотеки методів на базі платформи «.NET Standard 2.0». Для візуального відображення налаштування відповідних процесів розроблено віконний додаток на базі платформ «.NET 6.0» та «WPF». Для побудови необхідних графіків використано бібліотеку «OxyPlot». Досліджено якість роботи запропонованої згорткової нейронної мережі залежно від кількості блоків та розмірів фільтра згортки. Найкращі результати роботи досягаються із 3 блоками згорткових та агрегувального шарів та розмірами фільтра згортки 3 x 3 пікселів. Оптимальної точності класифікації об'єктів зображень мережа набуває при її навчанні протягом 14 epoch.

Ключові слова: нейрон; штучна нейронна мережа; зображення; класифікація об'єктів; згортка.

Актуальність теми. Розпізнавання та класифікація об'єктів зображень на сьогодні є розповсюдженою задачею, яка дедалі краще розв'язується завдяки використанню штучних нейронних мереж (ШНМ). Вони є одним із найбільш досліджуваних напрямів розробки штучного інтелекту та мають широкі можливості застосування в різноманітних галузях діяльності людини: медицині, економіці, обробці інформації, комп'ютерному зорі тощо. У задачах розпізнавання та класифікації об'єктів зображень потрібно оброблювати велику кількість даних, що зумовлює дві проблеми: по-перше, вибору оптимальної моделі нейронної мережі, по-друге, побудови ефективного алгоритму її навчання.

Аналіз останніх досліджень та публікацій. Суттєвим внеском у розвиток штучних нейронних мереж для класифікації об'єктів зображень стало створення проекту ImageNet. Він являє собою велику базу візуальних даних, що складається з більше ніж 14 мільйонів зображень, описаних вручну та призначених для дослідження програмного забезпечення, за допомогою якого здійснюється розпізнавання об'єктів на зображенні [1]. Також потрібно зазначити, що сучасні підходи до розпізнавання та класифікації об'єктів зображень зазвичай використовують згорткові нейронні мережі [2]. Високої продуктивності та точності в задачах класифікації об'єктів зображень із набору даних ImageNet можливо досягти із застосуванням поглибленої традиційної архітектури ConvNet [3], яка була запропонована ще в 1989 році для розв'язання задачі розпізнавання рукописних цифр [4]. Її суть полягає в використанні декількох блоків, що складаються з множини згорткових шарів із фільтром згортки невеликих розмірів та агрегувального шару.

Метою статті є дослідження використання згорткових нейронних мереж у задачах розпізнавання та класифікації об'єктів зображень, зокрема визначення статі за обличчям людини. Для проведення дослідження було розроблено програмний додаток для побудови згорткових нейронних мереж, їх навчання та тестування точності роботи.

Викладення основного матеріалу. Штучні нейронні мережі – це обчислювальні математичні моделі, які дозволяють із сукупності вхідних даних отримати необхідні вихідні дані. Вони імітують біологічні процеси, що відбуваються в нервовій системі людини.

Ключовими елементами нервової системи є сукупність нервових клітин, нейронів. Тіло нейрона називається сомою. У нейрона є відростки, якими в нього надходить інформація – дендрити, і відросток,

яким нейрон передає інформацію – аксон. Кожен нейрон може мати декілька дендритів і лише один аксон, що передає імпульс наступним нейронам нервової системи через нервові з'єднання – синапси [5].

Саме біологічні нейрони наштовхнули на створення штучних. Побудуємо спрощену модель штучного нейрона (рис. 1) та детальніше опишемо його роботу.

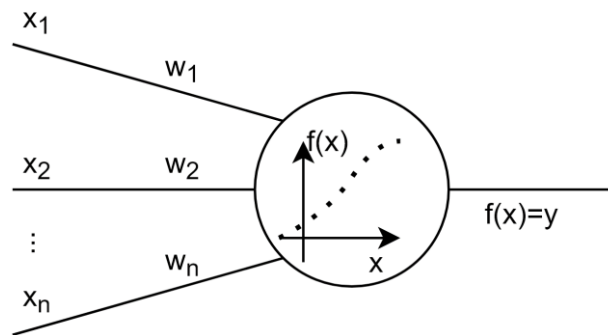


Рис. 1. Спрощена модель штучного нейрона

Зображений штучний нейрон (рис. 1) має n дендрити, синапси яких мають ваги $w_i, i = \overline{1, n}$, відповідно. $x_i, i = \overline{1, n}$, – це вхідні сигнали, які надходять до синапсів, $f(x)$ – деяка передатна функція або порогова активаційна функція, а y – вихідний сигнал. Надходячи до синапсів, вхідні сигнали масштабуються відповідно до їх вагових коефіцієнтів та подаються на вхід суматора, що оброблює їх за формулою:

$$x = \sum_{i=1}^n x_i w_i, \quad (1)$$

де x – аргумент порогової активаційної функції;

n – кількість дендритів;

x_i – вхідні сигнали;

w_i – вагові коефіцієнти синапсів.

Порогова активаційна функція має вигляд:

$$y = f(x), \quad (2)$$

де y – вихідний сигнал аксона нейрона.

Вибір функції активації залежить від необхідного вихідного сигналу та швидкодії ШНМ, а також визначається розв'язуваною задачею. Найбільш поширеною є сигмоїдальна функція, яка може бути дискретною або аналоговою [6]. Прикладом найпростішої активаційної функції є функція Гевісайда (рис. 2, а) – дискретний сигмоїд із параметром, який описується так:

$$y = \begin{cases} 1 & \text{при } x \leq a, \\ 0 & \text{при } x > a, \end{cases} \quad (3)$$

де a – деякий параметр.

Досить часто використовують сигмоїдальну логічну функцію з параметрами (рис. 2, б), яка є аналоговою, нелінійною, і вдало підходить до задач класифікації об'єктів зображень. Вона описується функцією:

$$y = \frac{b}{c + e^{dv}}, \quad (4)$$

де b, c, d – параметри функції.

Також досить поширеною є активаційна функція «ReLU» (rectified linear unit) (рис. 2, в), яка повертає вхідне значення у випадках, якщо воно додатне, і 0 – якщо від'ємне:

$$y = \begin{cases} 0 & \text{при } x \leq 0, \\ x & \text{при } x > 0. \end{cases} \quad (5)$$

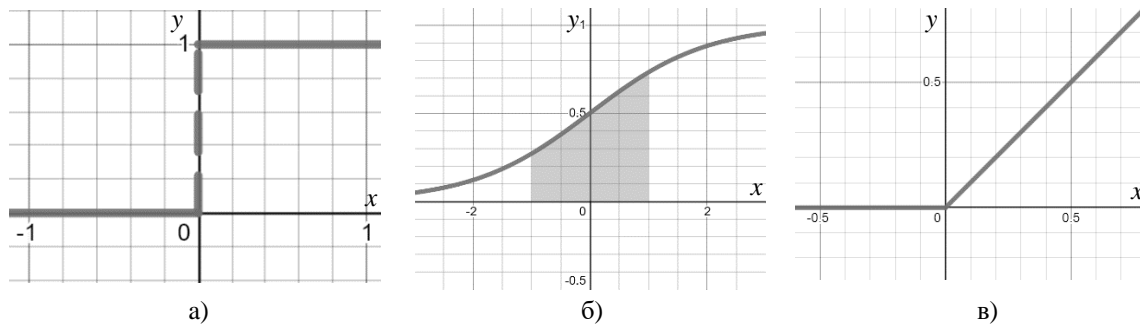


Рис. 2. Порогові активаційні функції: а – дискретний сигмоїд із параметром $a = 0$, б – сигмоїдальна логічна функція з параметрами: $b = 1$, $c = 1$, $d = -1$, в – функція «ReLU»

Існує велика кількість варіантів побудови штучних нейронних мереж. Вони відрізняються структурами – моделями зв'язків нейронів. Розрізняють ієрархічні, рекурентні та конкурентні структури. В цілому, усі моделі штучних нейронних мереж побудовані подібним чином. Нейрони розподіляються в мережі по шарах. Спочатку інформація надходить на нейрони вхідного шару. Вони передають її далі – до нейронів так званого прихованого шару. В прихованих шарах здійснюється основна обробка даних, після чого інформація надходить в останній шар – вихідний. Кількість нейронів та прихованих шарів обираються залежно від конкретної задачі, яку вирішує нейронна мережа, обсягу даних та доступних обчислюваних ресурсів.

Першою штучною нейронною мережею вважають перцептрон Розенблата (рис. 3), основою якого є модель роботи мозку при розпізнаванні оптичних образів [7]. Структурно – це проста рефлекторна мережа. За сучасною термінологією така ШНМ належить до одношарових із ієрархічною структурою. Входи такої ШНМ (S-елементи) моделюють роботу сітківки ока, що збирає інформацію про навколишнє середовище та передає її до А-елементів, які в свою чергу є насправді штучними нейронами з пороговою активаційною функцією. R-елементи слугують для формування реакції перцептрона, вони мають фіксовані вхідні ваги, які масштабуються для кожного А-елемента.

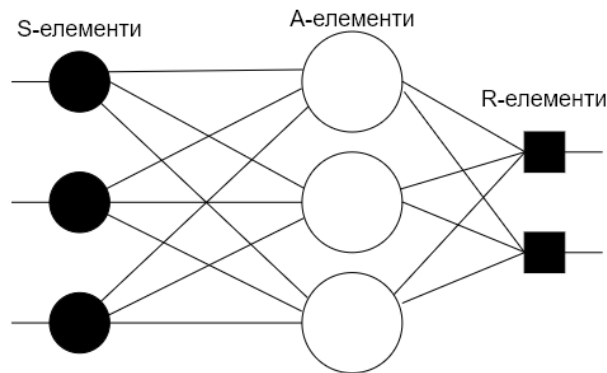


Рис. 3. Перцептрон Розенблата

При побудові нейронної мережі важливе значення має її навчання. Чим більшою буде кількість даних, що використовуються при навчанні, тим ефективнішим буде його результат. Дані можливо отримати ручним або автоматичним методом. Вручну зібрані дані, хоч і є точнішими, але, враховуючи необхідну їх кількість для якісного навчання ШНМ, зазвичай становлять непосильну задачу. Частіше використовують автоматичний збір даних. Досить розповсюдженою практикою є створення іншої ШНМ, яка підготовлюватиме дані та «змагатиметься» з першою. Правильний збір даних – це один із найскладніших та найоб'ємніших етапів для реалізації навчання нейронної мережі. Існують три основні підходи до навчання нейронної мережі: навчання із вчителем, навчання без вчителя та навчання з підкріпленням.

При навчанні із вчителем потрібно заздалегідь підготувати вибірку даних для навчання із множини пар, кожна з яких складається із даних, що надходять у вхідний шар ШНМ, та ідеального значення для вихідного шару – «правильної відповіді». Таке навчання спрямоване на досягнення такого стану нейронної мережі, щоби за заданих вхідних даних максимально наблизитися до ідеальних вихідних.

Для досягнення мети при навчанні без вчителя необхідно постійно модифікувати внутрішній стан нейронної мережі, тобто вагові коефіцієнти, згідно з оцінкою вхідних та набутих під час навчання даних. Бажані значення порівнюються з поточними та виконується корекція за допомогою зворотного зв'язку.

При навчанні з підкріпленням нейронної мережі не вказується точного значення бажаного виходу, але після отримання цього виходу їй виставляється оцінка – відповідний коефіцієнт відпрацювання ШНМ.

Для навчання досліджуваної ШНМ було використано метод зворотного поширення помилки – один із методів навчання зі вчителем. Ідея цього методу полягає в тому, що сигнал помилки поширюється від виходів ШНМ до її входів, тобто в зворотному напрямку. Його алгоритм являє собою ітеративний алгоритм градієнтного спуску, що використовується для мінімізації помилки роботи ШНМ шляхом пошарової корекції ваги її нейронів. На кожній ітерації вагові коефіцієнти ШНМ змінюються таким чином, аби покращити виконання поточного завдання.

Згорткові нейронні мережі (ЗНМ) – це ШНМ такої архітектури, яка побудована на основі багатшарового перцептрона, модифікованого так, аби підготовка вхідних даних була мінімальною. Такі нейронні мережі найчастіше застосовуються для розпізнавання та класифікації об'єктів зображень, у галузі комп'ютерного зору. Створення ЗНМ спонукалося біологічним процесом з'єднання нейронів зорової кори тварин. Нейрони кори головного мозку реагують на вхідний сигнал лише в області свого зорового поля, яке має назву рецептивного поля. Всі рецептивні поля можуть перекриватися один одним, але в результаті складаються так, щоби покрити всю область зорового поля. За такого підходу зображення оброблятимуться набагато швидше порівняно з класичними багатшаровими перцептронами. При цьому загальна структура у них схожа. Згорткова нейронна мережа також складається з шарів входу, виходу та декількох прихованих шарів, але в ЗНМ приховані шари містять згорткові, агрегувальні, повнозв'язні та шари нормалізації.

Згорткові шари є найбільш важливим елементом ЗНМ. Згорткові нейрони обробляють лише ті дані, які знаходяться в їх рецептивному полі [8].

Також важливими структурними елементами ЗНМ є агрегувальні шари. Вони можуть бути як глобальними – розповсюджуватися на окрему частину вхідних даних, так і локальними – діяти в рецептивних полях нейронів поточного шару. Головним завданням агрегувальних шарів є формування взаємного зв'язку між декількома нейронами попереднього шару і єдиним нейроном поточного шару. При цьому агрегувальний шар зменшує розмірність даних та зберігатиме найважливіші характеристики нейронів, які були знайдені за результатами роботи шару згортки. Через це при побудові ЗНМ, зазвичай, агрегувальні шари чергуються із шарами згортки.

Повнозв'язними називають такі шари, в яких всі нейрони попереднього шару пов'язані з нейронами наступного шару, що має місце в більшості класичних ШНМ, зокрема в перцептронах. Такі шари використовуються рідко, адже вони ускладнюють модель ЗНМ та інколи можуть призвести до ігнорування ознак, знайдених раніше за допомогою шарів згортки та агрегування.

У згортковій нейронній мережі кожен шар відповідає певній властивості зображення. В перших шарах містяться основні ознаки: контури, межі. Після їх обробки можливо розпізнати текстури, частини об'єктів. У вихідних шарах міститиметься результат класифікації зображення, виокремлений об'єкт або певна властивість зображення. Для розглядуваної задачі це буде результат класифікації об'єкта зображення: якщо обчислене вихідним нейроном значення є ближчим до 0, то нейронна мережа вирішила, що об'єктом зображення є жінка, інакше, якщо це значення є ближчим до 1, – чоловік.

Модель нейронної мережі, її шари, активаційні функції, алгоритм навчання потрібно обирати згідно з поставленою задачею. Для побудови пропонованої нейронної мережі було обрано класичну задачу – класифікацію об'єктів на зображенні. Почасти це зумовлено тим, що вже існує вдалий набір даних, який складається з 47009 зображень обличч чоловіків та жінок для навчання та 11649 зображень обличч для тестування ШНМ [9]. На виході нейронна мережа визначить обличчя чоловіка чи жінки є в завантаженому на її вході зображенні. Зважаючи на формат вхідних даних (масив значень пікселів), використовувати традиційні повнозв'язні нейронні мережі недоцільно через необхідність величезної кількості нейронів вхідного шару. Тут ЗНМ є найкращим вибором саме завдяки згортці, адже вона дозволяє зменшити обсяг інформації, що зберігається у пам'яті, а також ієрархічно виділити та агрегувати ознаки вхідних даних. Завдяки цьому збільшується обчислювальна потужність нейронної мережі та здатність моделювати більш складні залежності.

При побудові ШНМ було використано: декілька блоків згорткових та агрегувального шарів, шар вирівнювання, шари повнозв'язних нейронів, вихідний нейрон. Згорткові та агрегувальні шари потрібні для визначення загальних властивостей зображень. Шари повнозв'язних нейронів слугуватимуть для визначення ознак, притаманних обличчям саме чоловіків чи жінок. Вихідний шар – один, який становить повнозв'язний нейрон. Пороговою активаційною функцією для всіх нейронів, окрім вихідного, було обрано згадувану вище функцію «ReLU». Це дозволяє витрачати якомога менше ресурсів та не обраховувати значення менші за 0. Активаційною функцією вихідного нейрона є сигмоїдальна. Схематична модель такої нейронної мережі наведена на рисунку 4.

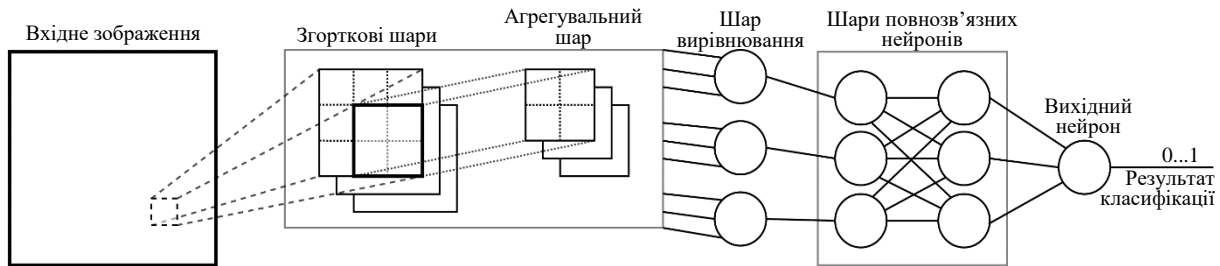


Рис. 4. Схематична модель розробленої ШНМ для класифікації об'єктів зображень

Для побудови запропонованої нейронної мережі, її навчання та тестування було обрано «TensorFlow». Це – бібліотека з відкритим кодом, яка призначена для побудови та навчання нейронних мереж, що розроблена, розвивається та підтримується компанією «Google». Для зручності розробки ШНМ додатково було використано високорівневий API «Keras.NET» [10]. Також для управління «Keras.NET» було вирішено розробити бібліотеку методів мовою програмування «C#» із використанням платформи розробки «.NET Standard 2.0».

Для візуального відображення налаштування процесів, які виконуватимуться в бібліотеці, згаданій вище, було розроблено віконний додаток також мовою програмування «C#» із використанням платформ розробки «.NET 6.0» та «WPF». Для побудови графіка залежності точності класифікації від епохи навчання використано бібліотеку «OxyPlot» [11].

Умовно додаток можна розділити на 3 частини: побудова ШНМ, її навчання, тестування.

При побудові нейронної мережі користувачу надається можливість обрати ширину та висоту, до якої буде стискатися зображення, та початкові параметри ШНМ: швидкість навчання та імпульс. Елементи згорткової нейронної мережі додаються блоками, які складаються зі згорткових та агрегувальних шарів. Надається можливість обирати кількість згорткових та повнозв'язних шарів, а також розміри згорткових та агрегувальних шарів. Як активаційну функцію можна обирати або «ReLU», або сигмоїдальну. Інтерфейс описаних можливостей додатка зображений на рисунку 5.

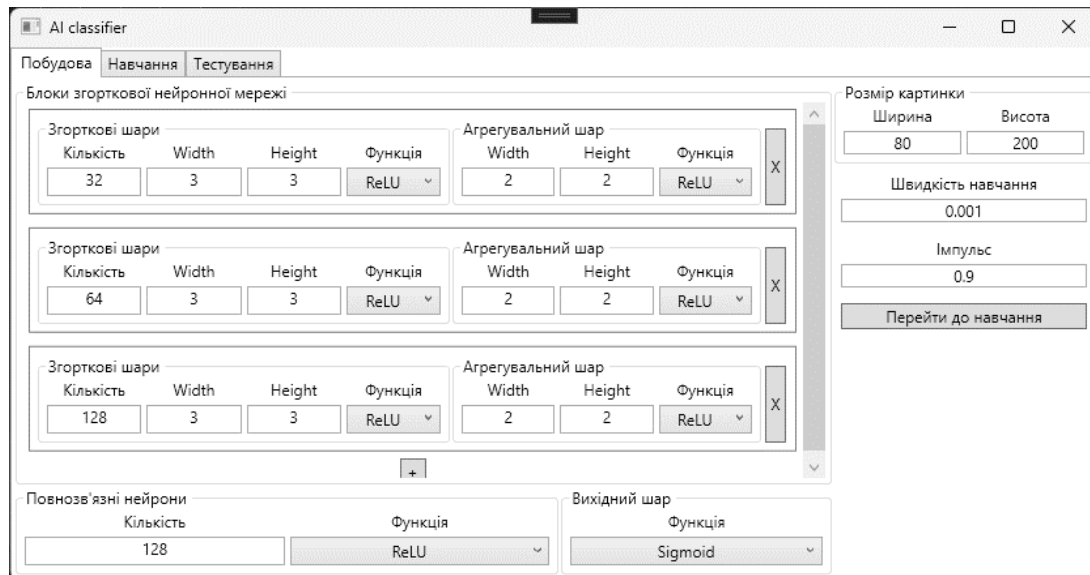


Рис. 5. Налаштування побудови штучної нейронної мережі

Наступною частиною є навчання нейронної мережі. Для цього спочатку потрібно обрати набір зображень, розділений на 2 теки, які іменовані об'єктами класифікації, тут це «чоловік» та «жінка». Для автоматичного тестування ШНМ необхідно обрати аналогічним чином розподілені дані. Далі потрібно вказати кількість епох та натиснути кнопку «Почати навчання». Під час процесу навчання виводиться графік залежності точності роботи ШНМ від епохи навчання. Отриману нейронну мережу можливо зберегти. Описаний інтерфейс зазначено на рисунку 6, а.

Для ручного тестування нейронної мережі першим кроком буде вибір зображення. На ньому виводитиметься прямокутник, який буде потрібно розташувати на обличчі людини. Можливо тестувати

як роботу щойно побудованої ШНМ, так і збереженої, завантаженої з відповідного файлу. Інтерфейс ручного тестування вказано на рисунку б, б.

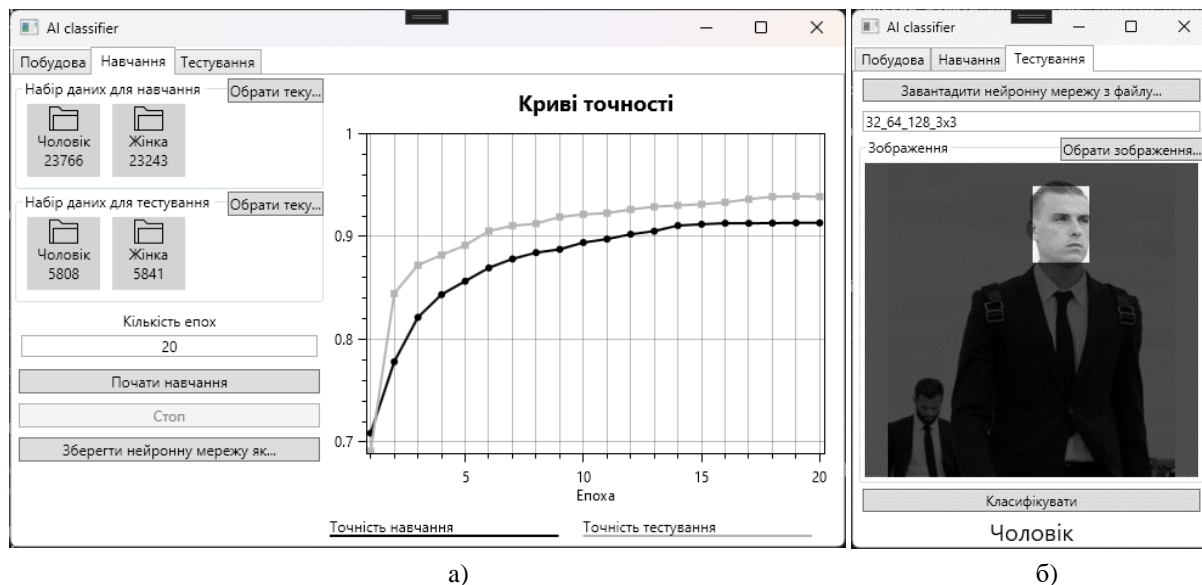


Рис. 6. Скріншоти додатка при роботі ШНМ із фільтром згортки розмірами 3 x 3 пікселів та 3 блоками згорткових і агрегуювального шарів: а – навчання ШНМ, б – ручне тестування

За допомогою створеного додатка було проведено дослідження якості ЗНМ залежно від кількості блоків та розмірів фільтра згортки (табл. 1). Кількість блоків змінювалася від 1 до 4, а розміри фільтра згортки – від 3 x 3 до 4 x 4 пікселів. Кількість епох навчання було обрано дещо надмірну – 20 епох.

Таблиця 1
Порівняння роботи ЗНМ з різною кількістю блоків та з різними розмірами фільтра згортки

Кількість блоків ЗНМ	Розміри фільтра згортки, пікселів	Точність навчання, %	Точність тестування, %
1	3 x 3	77,8	81,1
1	4 x 4	74,9	80,6
2	3 x 3	86,0	87,2
2	4 x 4	86,6	86,2
3	3 x 3	91,3	93,3
3	4 x 4	89,6	90,4
4	3 x 3	91,4	92,7
4	4 x 4	90,9	92,0

Дані було розподілено на 2 підмножини: одна використовувалася для навчання, друга – для тестування після кожної епохи. Варто зауважити, що дані для тестування не використовувалися впродовж навчання. Проаналізувавши отримані результати (табл. 1), можна дійти висновку, що збільшення розмірів фільтра згортки до 4 x 4 пікселів призводить до зменшення точності класифікації. Це відбувається через те, що збільшення розмірів фільтра згортки негативно впливає на коректність виокремлення певних властивостей зображення. Також збільшення розмірів фільтра згортки призводить до зменшення глибини ЗНМ, що в свою чергу зменшує точність класифікації. Для розглядуваної задачі обирати кількість блоків ЗНМ більшою за 3 недоцільно, оскільки вартість навчання мережі збільшується значною мірою, а точність класифікації – несуттєво.

Виходячи зі зроблених висновків, проаналізуємо графіки точності навчання та тестування нейронної мережі саме з 3 блоками згорткових та агрегуювального шарів та розмірами фільтра згортки 3 x 3 пікселів (рис. 6, а). Оптимальної точності класифікації об'єктів зображень ШНМ набуває при її навчанні протягом 14 епох. Збільшення кількості епох навчання є недоцільним.

Висновки та перспективи подальших досліджень. Пропонована згорткова нейронна мережа може бути вдало застосована в іншій ШНМ, що призначатиметься, наприклад, для виявлення та локалізації на

зображенні декількох об'єктів класифікації, для розпізнавання особи за обличчям або виразу обличчя. Також така ЗНМ може бути застосована в складі так званої генеративної змагальної ШНМ як дискримінатор, що досліджуватиметься в подальшому.

Список використаної літератури:

1. *Reynolds M.* New computer vision challenge wants to teach robots to see in 3D / *M.Reynolds* // *New Scientist*. – 2017 [Electronic resource]. – Access mode : <https://cutt.ly/sKlw30e>.
2. *Krizhevsky A.* ImageNet classification with deep convolutional neural networks / *A.Krizhevsky, I.Sutskever, E.Geoffrey* // *Communications of the ACM*. – 2017. – Vol. 6, № 60. – P. 84–90.
3. *Simonyan K.* Very Deep Convolutional Networks for Large-Scale Image Recognition / *K.Simonyan, A.Zisserman* // *International Conference on Learning Representations 2015*. – 2015 [Electronic resource]. – Access mode : <https://arxiv.org/pdf/1409.1556.pdf>.
4. *Handwritten Digit Recognition with a Back-Propagation Network* / *Y.Le Cun, B.Boser, J.Denker and other* // *Advances in Neural Information Processing Systems*. – 1989. – № 2. – P. 396–404.
5. *Neural Network Design* / *M.T. Hagan, H.B. Demuth, M.H. Beale, O.De Jesus*. – 2nd ed. – 2014. – 1012 p.
6. *Новотарський М.А.* Штучні нейронні мережі: обчислення / *М.А. Новотарський, Б.Б. Нестеренко*. – Київ : Інститут математики НАН України, 2004. – 408 с.
7. *Rosenblatt F.* The perceptron: a probabilistic model for information storage and organization in the brain / *F.Rosenblatt* // *Psychological Review*. – 1958. – № 65 (6). – P. 386–408.
8. *Goodfellow I.* *Deep Learning* / *I.Goodfellow, Y.Bengio, A.Courville*. – MIT Press, 2016. – 781 p.
9. *Chauhan A.* Gender Classification Dataset / *A.Chauhan*. – 2020 [Electronic resource]. – Access mode : <https://cutt.ly/NKleYG3>.
10. *Karra J.* Deep learning using Keras.Net with C# / *J.Karra*. – 2020 [Electronic resource]. – Access mode : <https://cutt.ly/4KleSIY>.
11. *Reznykov I.* Wpf application with real-time data in OxyPlot charts / *I.Reznykov*. – 2017 [Electronic resource]. – Access mode : <https://svitla.com/blog/wpf-application-with-real-time-data-in-oxyplot-charts>.

References:

1. Reynolds, M. (2017), «New computer vision challenge wants to teach robots to see in 3D», *New Scientist*, [Online], available at: <https://cutt.ly/sKlw30e>
2. Krizhevsky, A., Sutskever, I. and Hinton, E.G. (2017), «Advances in Neural Information Processing Systems», *Communications of the ACM*, Vol. 6, No. 60, pp.84–90.
3. Simonyan, K. and Zisserman, A. (2015), *Very Deep Convolutional Networks for Large-Scale Image Recognition*, *International Conference on Learning Representations 2015*, [Online], available at: <https://arxiv.org/pdf/1409.1556.pdf>
4. Le Cun, Y., Boser, B., Denker, J. et al. (1989), «Handwritten Digit Recognition with a Back-Propagation Network», *Advances in Neural Information Processing Systems*, No. 2, pp. 396–404.
5. Hagan, M.T., Demuth, H.B., Beale, M.H. and De Jesus, O. (2014), *Neural Network Design*, 2nd ed., 1012 p.
6. Novotars'kyj, M.A. and Nesterenko, B.B. (2004), *Shtuchni neyronni merezhi: obchyslennja*, Instytut matematyky NAN Ukraïny, Kyïv, 408 p.
7. Rosenblatt, F. (1958), «The perceptron: A probabilistic model for information storage and organization in the brain», *Psychological Review*, No. 65 (6), pp. 368–408.
8. Goodfellow, I., Bengio, Y. and Courville, A. (2016), *Deep Learning*, MIT Press, 781 p.
9. Chaunah, A. (2020), *Gender Classification Dataset*, [Online], available at: <https://cutt.ly/NKleYG3>
10. Karra, J. (2020), *Deep learning using Keras.Net with C#*, [Online], available at: <https://cutt.ly/4KleSIY>
11. Reznykov, I. (2017), *Wpf application with real-time data in OxyPlot charts*, [Online], available at: <https://svitla.com/blog/wpf-application-with-real-time-data-in-oxyplot-charts>

Кушнір Надія Олександрівна – асистент кафедри інженерії програмного забезпечення Державного університету «Житомирська політехніка».

<https://orcid.org/0000-0002-0797-3687>.

Наукові інтереси:

- комбінаторна оптимізація;
- інформаційні технології.

E-mail: kipz_kno@ztu.edu.ua.

Локтікова Тамара Миколаївна – старший викладач кафедри інженерії програмного забезпечення Державного університету «Житомирська політехніка».

<https://orcid.org/0000-0002-3525-0179>.

Наукові інтереси:

- цифрова обробка зображень;
- інформаційні системи та технології.

E-mail: tamlokt@ukr.net.

Морозов Андрій Васильович – кандидат технічних наук, доцент, проректор з науково-педагогічної роботи Державного університету «Житомирська політехніка».

<https://orcid.org/0000-0003-3167-0683>.

Наукові інтереси:

- комбінаторна оптимізація;
- інформаційні технології.

E-mail: morozov@ztu.edu.ua.

Юрченко Владислав Олегович – магістрант Державного університету «Житомирська політехніка».

<https://orcid.org/0000-0003-0608-6696>.

Наукові інтереси:

- інформаційні системи та технології.

E-mail: boo.yurchenko@gmail.com.

Kushnir N.O., Loktikova T.M., Morozov A.V., Yurchenko V.O.

Using convolutional neural networks in image object recognition and classification tasks

The usage of artificial neural networks in the tasks of object recognition and classification is studied. The classical problem of classifying objects in an image is considered, namely - determining the sex of a person by face. This is due to the fact that there is a successful set of data, which consists of 47,009 images of faces of men and women and 11,649 images of faces for training and testing artificial neural networks. It is suggested to use convolutional neural networks. This approach reduces the amount of information stored in memory, as well as hierarchically separates and aggregates features of the input data. Convolutional neural network consists of several blocks of convolutional and aggregating layers, alignment layer, layers of fully connected neurons, source neuron. The "ReLU" function was selected as the threshold activation function for all neurons except the original one. The activation function of the original neuron is sigmoidal. The neural network was built, trained and tested using the library "TensorFlow", API "Keras.NET", as well as the developed library of methods based on the platform ".NET Standart 2.0» and «WPF». The OxyPlot library was used to build the necessary graphs. The quality of work of the convolutional neural network that depends on the number of blocks and the sizes of the convolution filter has been investigated. The best results are achieved with 3 blocks of convolutional and aggregation layers and a convolutional filter size of 3x3 pixels. The optimum accuracy of image object classification is obtained by training the network for 14 epochs.

Keywords: neuron; artificial neural network; image; classification of objects; convolution.

Стаття надійшла до редакції 19.05.2022.