

Матвійчук А.М.

Державний університет інформаційно-комунікаційних технологій, Київ

АНАЛІЗ АРХІТЕКТУР НЕЙРОННИХ МЕРЕЖ ДЛЯ РОЗПІЗНАВАННЯ РУКОПИСНИХ ЦИФР

Анотація: Точність сучасних систем розпізнавання рукописних чисел демонструє значну варіативність, що залежить від технологій, обраних під час розробки конкретного методу. Нейронні мережі витіснили алгоритми, що покладаються на статистичні дані та набори даних для цього завдання, насамперед через їх кращу продуктивність.

Було розглянуто кілька останніх публікацій, присвячених проблемам розпізнавання рукописних цифр із застосуванням різних типів нейронних мереж. Ця стаття присвячена аналізу архітектури двох відомих типів: щільних нейронних мереж і згорткових нейронних мереж разом із відповідними методами навчання, застосованими до розпізнавання рукописних цифр із зображень.

Результати дослідження показують, що щільні нейронні мережі (DNN), незважаючи на те, що вони застосовні до поточного завдання, демонструють зниження продуктивності зі збільшенням розміру вхідного зображення. І навпаки, згорткові нейронні мережі (CNN) виявилися більш вправними в аналізі зображень завдяки своїм операціям згортки та об'єднання, таким чином встановлюючи їх пріоритет у цьому контексті.

Було досліджено чотири різні типи навчання, застосовні до нейронних мереж, а саме контрольоване навчання, неконтрольоване навчання, напівконтрольоване навчання та навчання з підкріпленням. Розслідування робить висновок, що навчання під контролем із використанням зворотного поширення є найбільш підходящим методом навчання нейронних мереж для вирішення виявленої проблеми. У статті далі досліджуються етапи та принципи, що лежать в основі роботи методу зворотного поширення.

Спираючись на отримані результати, існує план розробки програми, що використовує визначену архітектуру. Ця програма має на меті надати користувачам можливість ефективного вилучення цифрової інформації із зображень, незалежно від їх роздільної здатності. Передбачуваним результатом є програма, яка досягає високої точності та швидкості обробки при розпізнаванні та вилученні зображень.

Ключові слова: нейронні мережі, рукописні цифри, згорткові нейронні мережі, щільні нейронні мережі, алгоритми навчання нейронних мереж, метод зворотного поширення помилки.

Matviichuk A. M.

State University of Information and Communication Technology, Kyiv

ANALYSIS OF NEURAL NETWORK ARCHITECTURES FOR RECOGNITION OF HANDWRITTEN NUMBERS

Abstract: The accuracy of contemporary systems designed for the recognition of handwritten numbers exhibits significant variability, contingent upon the technologies selected during the development of a specific method. Neural networks have supplanted algorithms reliant on statistical data and datasets for this task, primarily due to their better performance.

Several recent publications addressing the challenges of handwritten digit recognition employing diverse types of neural networks have been examined. This paper delves into the analysis of the architecture of two prominent types: dense neural networks and convolutional neural networks,

along with their respective training methods applicable to the recognition of handwritten digits from images.

The study's findings reveal that Dense Neural Networks (DNN), while applicable to the task at hand, demonstrate diminished performance as the size of the input image increases. Conversely, Convolutional Neural Networks (CNN) prove more adept at image analysis, owing to their convolution and pooling operations, thus establishing their priority in this context.

Four distinct types of learning applicable to neural networks were explored, namely Supervised Learning, Unsupervised Learning, Semi-supervised Learning, and Reinforcement Learning. The investigation concludes that Supervised Learning using backpropagation emerges as the most fitting method for training neural networks to address the identified problem. The paper further explores the stages and principles underpinning the operation of the backpropagation method.

Building upon the derived results, there is a plan to develop an application leveraging the identified architecture. This application aims to empower users with the capability to efficiently extract digital information from images, irrespective of their resolution. The envisioned outcome is an application that attains high accuracy and processing speed in image recognition and extraction.

Keywords: neural networks, handwritten digits, convolutional neural networks, dense neural networks, neural network learning algorithms, backpropagation.

1. Вступ

Розпізнавання рукописного тексту – це процес перетворення тексту, створеного рукою людини (на папері, в різноманітних формах для малювання на комп'ютері тощо) у формат, який може бути зрозумілий обчислювальними системами (комп'ютер, смартфон, форма веб-сторінки тощо). Після написання тексту, його необхідно перевести у формат зображення (існують пристрої, що дозволяють виконувати цей процес прямо з паперу, але, як результат, система на вхід отримує оптичні дані) для подальшого аналізу програмним забезпеченням. Цей процес включає в себе виявлення символів на вихідному зображенні та їх класифікацію.

Однак, це зовсім не тривіальна задача, адже в процесі розпізнавання комп'ютер стикається з великим переліком складнощів, що сильно впливають на точність та швидкість цього процесу. Основними викликами є [1]:

1. Різноманітність почерку – кожна людина має унікальний стиль написання. Це означає, що один і той же символ може виглядати по-різному у відповідних почерках. Сюди також входить різноманітність способів написання символів (наприклад, літера Т може бути написана як «т» та «т»).
2. Наявність шуму – рукописний текст може бути зіпсований шумом на зображенні, таким як розмазування, вицвітання чорнила, сліди ручки тощо.
3. Зміна розміру та масштабу – рукописні символи можуть бути написані в різних розмірах та масштабах, що вимагає адаптації системи.
4. Перекриття символів – у рукописному тексті символи можуть перекриватися один одним, що робить їх виділення та розпізнавання важким завданням.
5. Обмежена якість зображення – якщо зображення рукописного тексту має низьку якість або низьку роздільну здатність, це може призвести до втрати важливої інформації і зробити розпізнавання менш точним.



Рис. 1. Шістнадцять способів написання цифри «2»

Зазначені складнощі відносяться як до всієї задачі розпізнавання рукописного тексту загалом, як і до розпізнавання чисел зокрема.

Технологічний процес, призначений для автоматизованого розпізнавання рукописного тексту називають Intelligent Character Recognition (ICR). Ця назва походить від попередника – OCR (Optical Character Recognition), що використовується для сканування та розпізнавання друкованого тексту.

2. Аналіз літературних даних і постановка проблеми

У публікації «Дослідження щодо використання алгоритмів на основі нейронних мереж для розпізнавання рукописних цифр» [2] автори досліджують науковий підхід в аналізі, використовуючи глибинне дослідження та мапування, для аналізу даних у галузі визнання рукописних цифр з використанням технік нейромереж. Комп'ютерний аналіз спрямований на визначення річного обсягу наукових досліджень щодо визнання рукописних цифр за допомогою нейронних мереж, темпів зростання та розподілу публікацій за установами. Представлено річний обсяг досліджень та шаблони зростання, демонструючи еволюцію досліджень з 2002 по 2018 рік. З цієї публікації можна зробити висновок, що нейронні мережі зробили великий внесок в точність розпізнавання рукописних чисел сучасних систем.

Точність сучасних систем розпізнавання рукописних чисел сильно варіюється в залежності від обраних технологій при розробці того чи іншого методу. ICR системи, що використовують алгоритми, статистичні дані та порівняння для визначення найбільшої кількості співпадінь мають точність, що не дозволяє їх практичне застосування [3]. Саме тому ця галузь перейшла до використання нейронних мереж. До того ж, окрім вищої точності, це дозволяє позбутися необхідності використання масивів даних при безпосередньому використанні, тому таким системам, зазвичай, не потрібне підключення до мережі Інтернет.

Стаття «Проблеми застосування штучних нейронних мереж в задачах розпізнавання образів» [1] присвячена докладному аналізу технічних викликів, що виникають у використанні штучних нейронних мереж (ШНМ) у контексті розпізнавання образів. У ній висвітлюється роль і важливість ШНМ як моделей для вирішення завдань розпізнавання образів та розкриваються основні питання, що ставляться перед вченими в цьому контексті. У розділі 4 цієї публікації автори фокусуються на технічних проблемах, що виникають у застосуванні ШНМ в задачах розпізнавання образів. Зокрема, детально розглядаються проблеми аналізу поведінки даних, нормалізації, масштабування, вивчення поведінки нейронів в прихованих шарах, а також висвітлюються виклики, пов'язані з великим обсягом обчислювальних витрат та тривалим часом навчання. Також, у статті визначають конкретні проблеми в застосуванні ШНМ до розпізнавання образів, такі як складність розпізнавання об'єктів з довільною орієнтацією, проблеми класифікації та локації, а також виклики, пов'язані з розпізнаванням образів у великих обсягах даних. Автори закликають до подальших досліджень та вдосконалень у вирішенні цих проблем для подальшого розвитку застосування ШНМ в області розпізнавання образів.

У статті авторства Елізабет Рані, Абхігна Редді та інших [4] досліджується застосування глибокої згорткової нейронної мережі (CNN) для розпізнавання рукописних чисел в банківській сфері. Основною метою дослідження є розробка та оптимізація архітектури CNN для точного та надійного визначення цифр, які можуть зустрічатися в банківських документах, зокрема в чеках та інших фінансових відомостях. Автори статті розглядають процес тренування глибокої згорткової нейронної мережі на наборі даних MNIST (Mixed National Institute of Standards and Technology). Дослідження також включає порівняльний аналіз ефективності згорткової нейронної мережі на різних етапах розпізнавання, зокрема етапу виділення ознак та етапу класифікації. Враховуються питання безпеки та швидкодії для забезпечення високої ефективності та надійності в застосуванні цієї технології у фінансовому

секторі. Головною ціллю дослідження автори зазначають бажання зробити банківські операції, де використовуються написані від руки числа, більш простими та безпомилковими.

Ще однією роботою, що використовує набір MNIST є «Розпізнавання рукописного тексту MNIST за допомогою згорткових нейронних мереж (CNN)» [5]. Автори представили офлайн метод розпізнавання рукописних цифр, що базується на різних техніках машинного навчання. Запропонована у статті техніка є більш корисною для непов'язаних рукописних даних у порівнянні з пов'язаними даними. Основна мета цієї статті – забезпечити ефективні та надійні методи ідентифікації транскрибованих цифр. Тобто, створена система не надто добре працює з зображеннями, на яких знаходиться більше однієї цифри. Не дивлячись на такий недолік, дана робота робить важливий внесок у розуміння процесу розробки ефективної нейронної мережі для розпізнавання рукописних чисел.

У роботі авторства Карен Симонян і Андрю Зиссермана [6], було представлено модель глибокої згорткової нейронної мережі, відому як VGG16, призначену для завдань класифікації великого обсягу зображень. Основною метою роботи було поліпшення точності класифікації за допомогою використання дуже глибокої архітектури.

Архітектура VGG16 складається з 16 шарів, що включають 13 блоків зі згортковими шарами та 3 повнозв'язаними шарами. Усі згорткові шари використовують ядра розміром 3x3, що дозволяє моделі вивчати більш складні функції зображень. Автори застосували фільтри з глибокою архітектурою, що дозволило їм здолати проблему зниклих градієнтів та покращити здатність моделі до вираження ієрархічних ознак.

3. Мета і задачі дослідження

Зростаюча кількість даних, що генеруються в даний час в різноманітних сферах життя ставить виклик у розробці більш точних та швидких методів розпізнавання рукописних чисел. З приходом цифровізації у наше життя, все більші об'єми даних опираються на послідовності цифр, адже саме в такому форматі комп'ютери мають найбільшу продуктивність для роботи з даними. Щоправда, суспільство та держави, що будувалися століттями, не мають змоги так швидко адаптуватися до сьогоденних реалій, тому й досі більшість документів, заявок, посвідчень тощо мають паперовий вигляд.

До прикладу, сучасний формат паспорта громадянина України (ID-картка) у порівнянні застарілим паспортом у формі книжечки має такі особливості:

- було додано унікальний номер запису в Єдиному державному демографічному реєстрі, що складається з 13 цифр, представлених двома послідовностями з восьми та п'яти цифр розділених текстовим символом «-» [7];
- серію та номер документу формату «EE000000» замінено на номер документу, що складається з дев'яти цифр;
- повну назву державного органу, яким видано документ було замінено на його номер, що складається з чотирьох цифр;
- було додано реєстраційний номер облікової картки платника податків (РНОКПП);
- назви місяців на документі замінені на їх порядковий номер у році.

Проте, Україна не була першою державою, що запровадила документи, з біометричними даними та переважанням цифрової інформації. Тенденція на такі зміни прийшла із Заходу, де цифровізація почала відбуватися раніше, тому ці речі стосуються переважної частини людства. Більшість цієї інформації є необхідною при заповненні паперових документів, бланків, заяв тощо. До того ж варто згадати про номер мобільного телефону, що за останні півтора десятки років став обов'язковим полем для заповнення в такого роду діяльності.

У разі необхідності перенесення заповнених документів в електронний формат, що, насправді, виконується у переважній більшості випадків, потрібно задіювати людський ресурс – людину, завдання якою вручну переписувати ці дані. Такий підхід, очевидно, не є

високопродуктивним, адже кожна літера та цифра займає відносно велику кількість часу, у порівнянні з тим, як з цією задачею міг впоратися комп'ютер.

Окрім документообігу, люди також займаються ручним сортуванням у сфері поштових перевезень. До прикладу, у разі надсилання листів або посилок, в яких дані, такі як поштовий індекс, номер будинку, квартири тощо, вказуються від руки. Хоча такий тип комунікацій здається застарілим та вже виходить з масового вжитку, він все ще являється важливим по всьому світу.

Також, покращення якості розпізнавання рукописних чисел може внести помітне покращення у галузь освіти та науки, а саме туди, де є математичні обрахунки. До прикладу, під час проведення аудиторних занять, або підготовці до них, часто є необхідним написання формул, розрахунків, математичних доведень тощо. У разі необхідності оцифрувати дані для створення звіту або презентації за темою занять, це може зайняти велику частину часу, адже наявні інструменти написання складних математичних обрахунків, через свою різноманітність та всеохопність, є нелегкими в освоєнні та навіть досвідчені користувачі можуть витратити більше часу, ніж така задача дійсно заслуговує.

Мета даної роботи полягає в пошуках оптимальної архітектури нейронної мережі для розпізнавання рукописних чисел.

Для досягнення поставленої мети необхідно вирішити такі завдання:

1. Проаналізувати види нейронних мереж для задачі класифікації зображень та обрати найефективнішу;
2. Проаналізувати архітектуру нейронної мережі;
3. Проаналізувати види навчання нейронних мереж.

4.1 Аналіз видів нейронних мереж

Нейронні мережі, є класом алгоритмів машинного навчання, що імітують структуру та функції біологічних нейронних мереж у мозку живих організмів. Ці моделі дозволяють автоматизоване вивчення та узагальнення складних залежностей у вхідних даних. Основними будівельними блоками нейронної мережі є штучні нейрони, що взаємодіють між собою через ваги та активаційні функції.

Основною характеристикою нейронної мережі є її архітектура, яка визначається кількістю шарів та кількістю нейронів у кожному з них. Першим шаром є вхідний, останнім — вихідний, а між ними можуть бути приховані шари. Кожен нейрон взаємодіє з нейронами попереднього та наступного шарів через ваги, які піддаються навчанню з метою оптимізації функції втрат.

Основним завданням нейронної мережі є мінімізація функції втрат, яка визначає різницю між передбачуваними та фактичними виходами. Цей процес досягається за допомогою алгоритмів оптимізації, таких як стохастичний градієнтний спуск, які підлаштовують ваги мережі для зменшення помилок прогнозування.

Сигнал передається через мережу від вхідного до вихідного шару, проходячи через кожен нейрон. Для введення нелінійності та здатності моделі вивчати складні залежності, застосовують функції активації. Популярні варіанти включають сигмоїдальні, гіперболічні та ReLU (Rectified Linear Unit) функції активації.

Види нейронних мереж, що можуть бути використані для реалізації заданої задачі:

- повністю зв'язана (щільна) нейронна мережа (Fully Connected Neural Network або Dense Neural Network)
- згорткова нейронна мережею (Convolution Neural Network, CNN).

Щільна нейронна мережа складається щонайменше з трьох шарів, в кожному з яких всі нейрони пов'язані з усіма нейронами попереднього та наступного рівня. Звідси і походить їх назва. Щільні нейронні мережі використовуються в різноманітних завданнях, включаючи класифікацію, регресію та розпізнавання об'єктів. Їхній широкий спектр застосувань пояснюється їхньою здатністю моделювати складні залежності в даних та ефективністю навчання на великих обсягах даних.

Проте, провівши дослідження виявилось, що щільні нейронні мережі не дуже добре підходять для розпізнавання об'єктів з зображень. Хоча вони можуть застосовуватися в таких цілях, вони не є найефективнішим засобом для таких задач з декількох причин:

1. зі збільшенням вхідного зображення їх розмірність швидко зростає, що призводить до зменшення продуктивності мережі та збільшення необхідних обчислювальних ресурсів (для зображення 28*28 пікселів кількість вхідних нейронів становить 784, для зображення 64*64 пікселя – 4096);
2. вони ігнорують просторові властивості зображень, такі як просторові локальність та інваріантність до трансформацій (зміна положення або кута нахилу об'єкта на зображенні сильно впливає на кінцевий результат);
3. щільні нейронні мережі можуть втрачати просторову інформацію зображення через агрегацію ваг (кожен нейрон отримує інформацію одразу з усіх пікселів, що негативно впливає на точність мережі).

Тому, було обрано застосовувати згорткову нейронну мережу. Згорткова нейронна мережа - це клас нейронних мереж, який виник для обробки та класифікації структурованих даних, таких як зображення. Вона використовує спеціальні шари, відомі як згорткові шари, для ефективного виявлення локальних особливостей та шаблонів у вхідних даних.

Зазвичай, згорткові нейронні мережі мають три види шарів:

1. Згорткові (convolutional), що виконують операції згортки вхідного зображення.
2. Об'єднуючі (pooling), що виконують операцію об'єднання сусідніх пікселів, вони можуть виконувати операції максимального, середнього або мінімального об'єднання;
3. Повністю зв'язані (dense) шари, що генерують результат на основі даних з попередніх шарів.

Зображення у таку нейронну мережу подається як набір матриць, що містять значення яскравості кожного пікселя у трьох кольорових каналах (червоний, зелений та блакитний, RGB) у вигляді чисел з рухомою комою.

Для проведення операції згортки, такі нейронні мережі використовують фільтри (ядра), що являють собою матриці значень з рухомою комою. Ця операція дозволяє виділити такі ознаки, як границі, текстури та структури. Кількість та розмірність фільтрів задається при створенні нейронної мережі. Згортка виконується кожним каналом для кожного фільтру, що в результаті дає набір матриць, кількість яких дорівнює кількості фільтрів, розмірністю:

$$N - M + 1$$

де N – довжина сторони вхідної матриці,

M – довжина сторони фільтру згорткового шару.

Операція згортки визначається наступним чином для двовимірної матриці (зображення) I та ядра K [8]:

$$S(i, j) = (I * K)(i, j) = \sum_m \sum_n I(m, n) K(i - m, j - n)$$

де $S(i, j)$ - значення пікселя результуючої матриці, $I(m, n)$ - значення пікселя вхідного зображення, $K(i-m, j-n)$ - значення ядра. Також, після операції згортки застосовується операція додавання зсуву (bias), що є важливим компонентом для забезпечення гнучкості та адаптивності мережі до різних видів даних. Після виконання всіх вищезазначених дій, застосовується функція активації нейронів, що для згорткового шару зазвичай є ReLU.

Після виконання згортки виконується операція об'єднання значень для кожної матриці (карт ознак) з попереднього шару. Вона використовуються для зменшення просторового розміру карт ознак, зберігаючи при цьому важливі інформаційні характеристики, що в свою чергу сприяє поліпшенню обчислювальної ефективності мережі. Найпоширенішими операціями об'єднання є максимальне та середнє, що використовують неперекриваючі області для взяття пікселів з регіонів зображення. Для створення нейронної мережі для розпізнавання рукописних чисел з чорно-білими зображеннями краще підходить операція максимального об'єднання, що визначається наступним чином:

$$P(i, j) = \max_{m, n} (I(i * s + m, j * s + n))$$

де $P(i, j)$ – значення пікселя у результуючій матриці після максимального об'єднання,
 $I(i * s + m, j * s + n)$ – значення пікселя вхідної матриці,
 s – крок об'єднання (stride),
 m та n – розмірність вікна об'єднання.

Параметри s , m та n задаються при створенні нейронної мережі. Операції об'єднуючого шару не потребують функцій активації, зсуву та навчання.

Після виконання операцій згортки та об'єднання набір карт ознак переходить для генерації результату на шар повністю зв'язаних нейронів, що має розмірність

$$N * M * K$$

де N – кількість карт ознак,
 M та K – їх розмірності.

Оскільки нейрони цього шару не приймають дані у вигляді масивів, перед передачею даних до набору застосовується операція розгортки в одновимірний масив (вектор). Ця операція часто використовується у контексті обробки даних у машинному навчанні або обчисленнях, де потрібно перетворити багатовимірні структури даних у вектор для подальших операцій.

Оскільки щільний шар нейронної мережі є вихідним, кількість вихідних нейронів визначається кількістю можливих відповідей – 10. Визначена цифра позначатиметься індексом найбільшого елементу масиву. Для кожного вихідного нейрону значення вагової суми визначається за формулою [9]:

$$z_i = \sum_j w_{ij}^{(L)} \cdot a_j^{(L-1)} + b_i^{(L)},$$

де L вказує на останній (вихідний) шар мережі,

$w_{ij}^{(L)}$ - вага між j -м нейроном у попередньому шарі та i -м нейроном у вихідному шарі,

$a_j^{(L-1)}$ - вихід j -го нейрона у попередньому шарі,

$b_i^{(L)}$ - зсув для i -го нейрона у вихідному шарі.

Найкращою функцією активації для задач класифікації є softmax, або ж нормована експоненційна функція [10], що в даному контексті має формулу:

$$y_i = \frac{e^{z_i}}{\sum_k e^{z_k}},$$

де y_i – вихід i -го нейрона після застосування функції,

e – число Ейлера,

z_i – вагова сума для i -го нейрона,

k – кількість класів, для яких визначаються ймовірності в рамках функції softmax.

4.2 Навчання нейронної мережі

Перед використанням нейронної мережі її необхідно буде навчити на наявному наборі даних. Існують декілька видів навчання нейронних мереж:

1. Навчання з учителем (Supervised Learning):

- У цьому виді навчання для кожного вхідного прикладу надається пара "вхід-вихід" (input-output), і мережа навчається знаходити відповідність між входом і вихідним значенням.
- Головна мета - мінімізація різниці між передбаченими вихідними значеннями і справжніми вихідними значеннями.
- Приклади: класифікація, регресія.

2. Навчання без учителя (Unsupervised Learning):

- У цьому виді навчання модель отримує тільки вхідні дані і намагається виявити приховані структури чи шаблони в цих даних.

- Головна мета - групування, розкриття прихованих залежностей або побудова представлення даних.
 - Приклади: кластеризація, автокодування.
3. Навчання з частковим навчанням (Semi-supervised Learning):
- Комбінує елементи навчання з учителем та навчання без учителя.
 - Має невелику кількість прикладів з відомими вихідними значеннями і багато прикладів без них.
 - Мета - використовувати велику кількість невідомих даних для поліпшення навчання.
 - Приклади: класифікація з обмеженим набором міток.
4. Навчання з підкріпленням (Reinforcement Learning):
- Модель (агент) взаємодіє з оточенням і отримує винагороду чи покарання в залежності від своїх дій.
 - Мета - вивчення оптимальної стратегії для максимізації накопиченої винагороди в часі.
 - Використовується в задачах прийняття рішень та управління.

Оскільки розпізнавання рукописних чисел є задачею класифікації, найкращим видом навчання буде навчання з учителем. Так, при передачі зображення при навчанні мережі, необхідно буде передавати очікуваний результат від нейронної мережі.

Для реалізації такого навчання необхідно реалізувати метод зворотного поширення помилки (backpropagation) для мінімізації функції втрат шляхом оновлення ваг та зсувів у нейронній мережі.

Процес backpropagation відбувається у кілька етапів [11]:

1. Прямий прохід (Forward Pass):
 - Вхідні дані подаються в мережу, і вона виконує прямий прохід, обчислюючи вихідні значення для кожного шару від входу до виходу.
2. Обчислення функції втрат:
 - Порівнюються передбачені вихідні значення з справжніми, і обчислюється функція втрат, яка відображає різницю між ними.
3. Зворотне поширення помилки (Backward Pass):
 - Розпочинаючи з останнього шару, обчислюється градієнт функції втрат відносно ваг та зсувів.
 - Градієнт передається в зворотного напрямку через мережу, використовуючи правило ланцюга (chain rule).
4. Оновлення параметрів:
 - Оновлення ваг та зсувів відбувається в напрямку, протилежному градієнту, з використанням методу градієнтного спуску.
 - Коефіцієнт швидкості навчання визначає, наскільки швидко оновлюються параметри.

Цей ітеративний процес повторюється протягом кількох епох або до досягнення задовільної збіжності. Метод зворотного поширення помилки дозволяє мережі автоматично коригувати свої параметри, адаптуючись до особливостей вхідних даних і покращуючи точність передбачень.

5. Висновки і перспективи подальших досліджень

У даній статті було розглянуто два види нейронних мереж, що можуть бути використані для задачі класифікації зображень за рукописними цифрами, що вони містять. За результатами дослідження, було виявлено, що щільні нейронні мережі (Dense Neural Networks, DNN), хоча й можуть бути застосовані для такої задачі, стають низькопродуктивними зі збільшенням розміру вхідного зображення. Згорткові нейронні мережі (Convolutional Neural Networks,

CNN) виявилися більш пристосованими для аналізу зображень, тому їх використання є пріоритетним.

Було детально розглянуто архітектуру та аспекти роботи згорткових нейронних мереж. Обрано вид навчання нейронної мережі, що найкраще підходить для вирішення поставленої задачі. Розглянуто етапи та принцип роботи методу зворотного поширення помилки (backpropagation).

На основі даного дослідження планується реалізувати додаток, що використовуючи обрану архітектуру дозволить користувачам виділяти цифрову інформацію з зображень з високою точністю та швидкістю, незалежно від роздільної здатності зображень.

Список використаної літератури:

1. Comprehensive review of artificial neural network applications to pattern recognition [Електронний ресурс] / Oludare Isaac Abiodun [та ін.] // IEEE access. – 2019. – Т. 7. – С. 158820–158846.
2. A survey on using neural network based algorithms for hand written digit recognition [Електронний ресурс] / Muhammad Ramzan [та ін.] // International journal of advanced computer science and applications. – 2018. – Т. 9, № 9.
3. Al-Taee M. M. Handwritten Recognition: a survey [Електронний ресурс] / May Mowaffaq Al-Taee, Sonia Ben Hassen Neji, Mondher Frikha // 2020 IEEE 4th international conference on image processing, applications and systems (IPAS), Genova, Italy, 9–11 груд. 2020 р. – [Б. м.], 2020.
4. MNIST handwritten digit recognition using machine learning [Електронний ресурс] / Elizabeth Rani G [та ін.] // 2022 2nd international conference on advance computing and innovative technologies in engineering (ICACITE), Greater Noida, India, 28–29 квіт. 2022 р. – [Б. м.], 2022
5. Digit recognition of MNIST handwritten using convolutional neural networks (CNN) [Електронний ресурс] / Sakthimohan M [та ін.] // 2023 international conference on intelligent systems for communication, iot and security (iciscois), Coimbatore, India, 9–11 лют. 2023 р. – [Б. м.], 2023.
6. Simonyan K. Very Deep Convolutional Networks for Large-Scale Image Recognition / Karen Simonyan, Andrew Zisserman // ICLR 2015, Сан-Дієго, 7 лип. 2014 р. – [Б. м.].
7. Що таке УНЗР та де його взяти? [Електронний ресурс] // Державна міграційна служба.
8. Dumoulin V. A guide to convolution arithmetic for deep learning / Vincent Dumoulin, Francesco Visin. – [Б. м.] : Монреа. ун-т, 2018. – 31 с.
9. LeCun Y. Deep learning [Електронний ресурс] / Yann LeCun, Yoshua Bengio, Geoffrey Hinton // Nature. – 2015. – Т. 521, № 7553. – С. 436–444.
10. Improved handwritten digit recognition using convolutional neural networks (CNN) [Електронний ресурс] / Savita Ahlawat [та ін.] // Sensors. – 2020. – Т. 20, № 12. – С. 3344.
11. Zhang D. The application research of neural network and BP algorithm in stock price pattern classification and prediction [Електронний ресурс] / Dehua Zhang, Sha Lou // Future generation computer systems. – 2021. – Т. 115. – С. 872–879.

References:

1. Comprehensive review of artificial neural network applications to pattern recognition [Electronic resource] / Oludare Isaac Abiodun [et al.] // IEEE access. – 2019. – Vol. 7. – P. 158820–158846.
2. A survey on using neural network based algorithms for hand written digit recognition [Electronic resource] / Muhammad Ramzan [et al.] // International journal of advanced computer

science and applications. – 2018. – Vol. 9, No. 9.

3. Al-Tae M. M. Handwritten Recognition: a survey [Electronic resource] / May Mowaffaq Al-Tae, Sonia Ben Hassen Neji, Mondher Frikha // 2020 IEEE 4th international conference on image processing, applications and systems (IPAS), Genova, Italy, December 9–11 2020 - [B. m.], 2020.

4. MNIST handwritten digit recognition using machine learning [Electronic resource] / Elizabeth Rani G [et al.] // 2022 2nd international conference on advance computing and innovative technologies in engineering (ICACITE), Greater Noida, India, April 28–29. 2022 - [B. m.], 2022.

5. Digit recognition of MNIST handwritten using convolutional neural networks (CNN) [Electronic resource] / Sakthimohan M [et al.] // 2023 international conference on intelligent systems for communication, iot and security (iciscois), Coimbatore, India, 9– February 11 2023 - [B. m.], 2023.

6. Simonyan K. Very Deep Convolutional Networks for Large-Scale Image Recognition / Karen Simonyan, Andrew Zisserman // ICLR 2015, San Diego, July 7. 2014 - [B. m.].

7. What is UNZR and where to get it? [Electronic resource] // State Migration Service.

8. Dumoulin V. A guide to convolution arithmetic for deep learning / Vincent Dumoulin, Francesco Visin. - [B. m.] : Monrea. University, 2018. - 31 p.

9. LeCun Y. Deep learning [Electronic resource] / Yann LeCun, Yoshua Bengio, Geoffrey Hinton // Nature. – 2015. – Vol. 521, No. 7553. – P. 436–444.

10. Improved handwritten digit recognition using convolutional neural networks (CNN) [Electronic resource] / Savita Ahlawat [et al.] // Sensors. – 2020. – Vol. 20, No. 12. – P. 3344

11. Zhang D. The application research of neural network and BP algorithm in stock price pattern classification and prediction [Electronic resource] / Dehua Zhang, Sha Lou // Future generation computer systems. – 2021. – Vol. 115. – P. 872–879.