

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет комп'ютерної інженерії та управління
(повна назва)

Кафедра електронних обчислювальних машин
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

Рівень вищої освіти перший (бакалаврський)

Використання згорткових нейронних мереж
для розв'язання задачі розпізнавання
реєстраційних автомобільних номерів
(тема)

Виконав:
здобувач 4 року навчання,
групи КІУКІ-21-5
Валентин ДЕМЧЕНКО
(власне ім'я, прізвище)

Спеціальність 123 «Комп'ютерна інженерія»
(код і повна назва спеціальності)

Тип програми освітньо-професійна
(освітньо-професійна або освітньо-наукова)

Освітня програма Комп'ютерна інженерія
(повна назва освітньої програми)

Керівник: доц. Наталія БОЛОГОВА
(посада, власне ім'я, прізвище)

Допускається до захисту

Завідувач кафедри ЕОМ Андрій КОВАЛЕНКО
(підпис) (власне ім'я, прізвище)

2025 р.

Харківський національний університет радіоелектроніки

Факультет _____ комп'ютерної інженерії та управління _____

Кафедра _____ електронних обчислювальних машин _____

Рівень вищої освіти _____ перший (бакалаврський) _____

Спеціальність _____ 123 «Комп'ютерна інженерія» _____
(код і повна назва)

Тип програми _____ освітньо-професійна _____
(освітньо-професійна або освітньо-наукова)

Освітня програма _____ Комп'ютерна інженерія _____
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

“ _____ ” _____ 20__ р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві _____ Демченку Валентину Павловичу _____
(прізвище, ім'я, по батькові)

1. Тема роботи Використання згорткових нейронних мереж для розв'язання задачі розпізнавання реєстраційних автомобільних номерів

затверджена наказом по університету від “ 26 ” травня 2025 р. № 424 Ст

2. Термін подання здобувачем роботи до екзаменаційної комісії _____ 16 червня 2025 р.

3. Вхідні дані до роботи 1) модель нейронної мережі: архітектура LeNet; 2) бібліотека: Keras; 3) датасет: Train_Cells із ресурсу Kaggle; 4) мова програмування: Python.

4. Перелік питань, що потрібно опрацювати у роботі _____

1) аналіз проблеми та огляд існуючих рішень;

2) аналіз проблем розпізнавання графічних образів;

3) аналіз методів розпізнавання графічних образів;

4) розробка рішення реалізації проєкту;

5) програмна реалізація алгоритму;

6) висновки.

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій – 13 слайдів.

6. Консультанти розділів роботи (заповнюється за наявності консультантів згідно з наказом, зазначеним у п.1)

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Строк / терміни виконання етапів роботи	Примітка
1	Аналіз проблеми та огляд існуючих рішень	27.05.25-30.05.25	
2	Аналіз проблеми розпізнавання графічних образів	31.05.25-03.06.25	
3	Розробка рішення реалізації проекту	04.06.25-07.06.25	
4	Програмна реалізація алгоритму	08.06.25-09.06.25	
5	Оформлення матеріалів кваліфікаційної роботи	10.06.25-11.06.25	
6	Подання кваліфікаційної роботи керівникові та її попередній захист	12.06.25-13.06.25	
7	Подання кваліфікаційної роботи на рецензування	14.06.25-16.06.25	

Дата видачі завдання “ 26 ” травня 2025 р.

Здобувач

_____ (підпис)

Керівник роботи

_____ (підпис)

доц. Наталія БОЛОГОВА

_____ (посада, власне ім'я, прізвище)

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: 54 с., 10 рис., 2 дод., 14 джерел.

ЗГОРТКОВА НЕЙРОННА МЕРЕЖА, КОМП'ЮТЕРНИЙ ЗІР, РОЗПІЗНАВАННЯ ГРАФІЧНИХ ОБРАЗІВ, СЕГМЕНТАЦІЯ ЗОБРАЖЕННЯ, РЕЄСТРАЦІЙНИЙ АВТОМОБІЛЬНИЙ НОМЕР, ДЕРЖАВНИЙ НОМЕРНИЙ ЗНАК.

Метою кваліфікаційної роботи є вивчення технологій комп'ютерного зору та написання програми, що здійснює сегментацію та розпізнавання символів реєстраційного автомобільного номера на зображенні.

В результаті роботи було проаналізовано існуючі методи розпізнавання графічних образів, розроблено алгоритм сегментації та розпізнавання букв і цифр державного номерного знака транспортного засобу на зображенні з використанням згорткової нейронної мережі. У процесі дослідження було вивчено основні етапи обробки зображень, включаючи попередню обробку, виділення контурів та нормалізацію символів.

Особливу увагу приділено формуванню навчальної вибірки та підбору архітектури згорткової нейронної мережі для досягнення високої точності розпізнавання. Були проведені експерименти з різними підходами до сегментації номерного знака на зображенні з урахуванням впливу зовнішніх факторів, таких як освітлення та якість знімка. Результати тестування продемонстрували ефективність розробленого методу при обробці зображень в реальному часі.

Отриманий програмний продукт може бути використаний у системах автоматизованого контролю транспорту та відеоспостереження.

ABSTRACT

Bachelor's thesis: 54 pages, 10 figures, 2 appendices, 14 sources.

CONVOLUTIONAL NEURAL NETWORK, COMPUTER VISION, GRAPHIC IMAGE RECOGNITION, IMAGE SEGMENTATION, REGISTRATION NUMBER PLATE, STATE NUMBER PLATE.

The major goal of this thesis is the qualification work is to study computer vision technologies and write a program that performs segmentation and recognition of characters in a vehicle registration number on an image.

As a result of the work, existing methods of graphic image recognition were analyzed, and an algorithm for segmentation and recognition of letters and numbers of a vehicle's license plate in an image using a convolutional neural network was developed. The research examined the main stages of image processing, including preprocessing, contour extraction, and symbol normalization.

Particular attention was paid to the formation of a training sample and the selection of a convolutional neural network architecture to achieve high recognition accuracy.

Experiments were conducted with different approaches to license plate segmentation in images, taking into account the influence of external factors such as lighting and image quality. The test results demonstrated the effectiveness of the developed method in real-time image processing.

The resulting software product can be used in automated traffic control and video surveillance systems.

ЗМІСТ

СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ	7
ВСТУП	8
1 ТЕХНОЛОГІЇ КОМП'ЮТЕРНОГО ЗОРУ	10
1.1 Теоретичні аспекти	10
1.2 Існуючі системи розпізнавання	12
1.3 Проблеми систем розпізнавання	14
2 МЕТОДИ РОЗПІЗНАВАННЯ ГРАФІЧНИХ ОБРАЗІВ	16
2.1 Основні методи розпізнавання образів	16
2.2 Розпізнавання за ознаками	17
2.3 Структурні методи розпізнавання	18
2.4 Статистичні методи.....	19
2.5 Штучні нейронні мережі	20
2.5.1 Концептуальна модель	20
2.5.2 Функції активації нейронів	26
2.5.3 LeNet.....	28
3.ПРОГРАМНА РЕАЛІЗАЦІЯ АЛГОРИТМУ	32
3.1 Постановка задачі.....	32
3.2 Засоби розробки рішення	33
3.3 Набір даних для навчання НМ.....	34
3.4 Структура алгоритму	36
3.5 Сегментація зображень.....	37
3.6 Архітектура нейронної мережі	38
3.7 Приклад роботи алгоритму	40
ВИСНОВКИ.....	42
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ	43
ДОДАТОК А Графічний матеріал кваліфікаційної роботи.....	45
ДОДАТОК Б Вихідний код	52

СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ

ДНЗ – дорожній номерний знак

ЗНМ – згорткові нейронні мережі

НМ – нейронні мережі

ТЗ – транспортний засіб

ШНМ – штучні нейронні мережі

ВСТУП

В даний час відбувається бурхливий розвиток інформаційних технологій. З їх допомогою автоматизуються багато процесів в самих різних галузях – в промисловості, бізнесі, медицині, освіті, транспорті та побутовій сфері. В області забезпечення безпеки та контролю дорожнього руху теж є проблеми, вирішення яких можливе завдяки застосуванню комп'ютерних технологій.

У сучасному світі наявність автомобіля є потребою мало не кожної людини. Число автомобілів на дорогах стрімко зростає.

За даними Укравтопром, продажі нових легковиків в Україні у 2024 році оцінено у понад 125 млрд гривень, це на 18% більше, ніж у 2023 році [1]. Понад три чверті від всіх легковиків, що у 2023 році поповнили вітчизняний автопарк були вживані авто, ввезені з-за кордону, за 12 місяців 2024 року українці придбали понад 222,1 тис. таких автомобілів [2].

Транспортна інфраструктура розвивається швидкими темпами, і необхідність контролювати рух транспорту зростає.

Розпізнавання реєстраційних номерних знаків – це одна із задач ідентифікації транспортних засобів (ТЗ). Системи розпізнавання реєстраційних номерів по зображенню і відео використовуються в багатьох сферах, пов'язаних з транспортом, наприклад, на автобазах, підприємствах, стоянках ТЗ, автозаправних станціях, станціях технічного обслуговування, біля терміналів платних доріг, при автоматичній фото- і відеофіксації порушень на трасах і т. д.

Використання систем ідентифікації транспорту значно збільшує швидкість роботи багатьох підприємств, дозволяє контролювати рух ТЗ, підвищує рівень безпеки дорожнього руху, запобігає злочинності.

На даний момент існує декілька алгоритмів розпізнавання автомобільних номерів, розроблені апаратні пристрої для використання цих

алгоритмів, але через високу вартість таких продуктів і складність роботи поки не відбулося їх масове впровадження [3]. Тому актуальні задачі розробки нових алгоритмів розпізнавання графічних образів і оптимізації існуючих рішень.

1 ТЕХНОЛОГІЇ КОМП'ЮТЕРНОГО ЗОРУ

1.1 Теоретичні аспекти

Основи комп'ютерного зору – яке також називають технічним зором і машинним зором – були закладені ще в середині двадцятого століття. Тим не менш, ця область науки і техніки вважається досить новою і активно розвивається в даний час.

Комп'ютерний зір (Computer Vision) тісно пов'язаний з машинним навчанням, комп'ютерною графікою і обробкою графічної інформації. Його головним завданням є ідентифікація об'єктів на зображеннях і на відеопотоці, тобто розпізнавання образів.

Системи комп'ютерного зору являють собою сукупність технологій і алгоритмів, за допомогою яких комп'ютер обробляє зображення і відеофайли, ідентифікує об'єкти на зображеннях, класифікує і аналізує їх. Такі системи отримують зображення і відеофайли за допомогою фотокамер, відеокамер, сканерів, рендерингу 2D- і 3D-моделей і т. д.

Загальну схему вирішення задачі машинного зору можна побачити на рисунку 1.1.

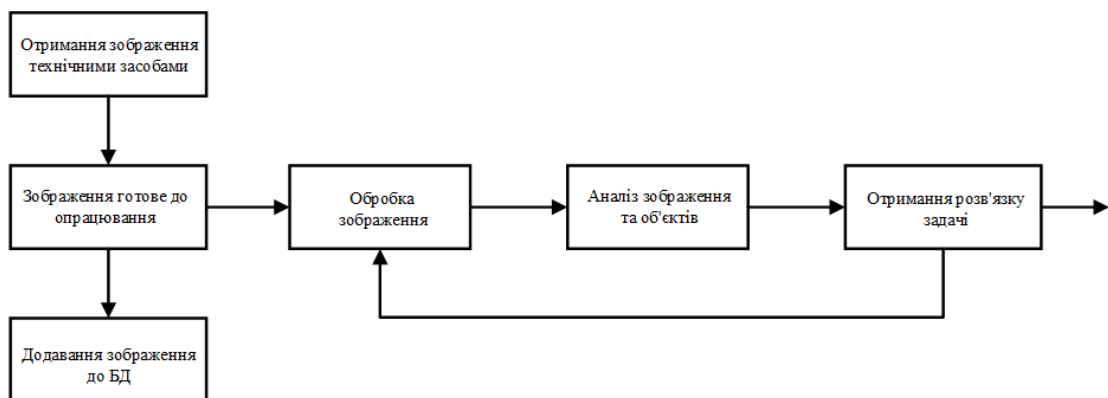


Рисунок 1.1 – Схема вирішення задачі машинного зору

В залежності від сфери діяльності в блоці схеми «Обробка зображення» будуть присутні різні алгоритми. Вибір методів і технологій обробки і аналізу графічної інформації залежить від розміру і якості зображень, ракурсів зйомки, кількості об'єктів для виявлення, наявності або відсутності бази даних і її обсягу [4]. У найпростішому випадку рішення задачі може мати двозначну відповідь (наприклад, «так, ні»), а в більш складному – вивід розпізнаного об'єкта або множини об'єктів, класифікація об'єктів, їх пошук в базі даних та інше.

Одним з найбільш поширених класів алгоритмів для розробки систем комп'ютерного зору є штучні нейронні мережі (ШНМ), які розглянуті в розділі 2.

Для роботи системи машинного навчання, заснованої на технології нейронних мереж, потрібне досить потужне обладнання. Організація роботи великих нейронних мереж обходиться дорого не тільки з точки зору бюджету, але і з точки зору обчислювальних ресурсів. У сучасних умовах активно розробляються апаратні рішення, здатні забезпечувати ефективне функціонування нейронних мереж на компактному обладнанні. Одним із таких є модуль Intel Movidius Myriad X MX HDDL M.2 – апаратний прискорювач з форм-фактором M.2, який інтегрується безпосередньо в системну плату або вбудовані пристрої. Цей модуль оснащений Vision Processing Unit (VPU) Intel Movidius Myriad X, що забезпечує високопродуктивне та енергоефективне виконання глибоких згорткових нейронних мереж (CNN) для задач класифікації зображень і розпізнавання об'єктів.

Intel MX HDDL M.2 підтримує апаратне прискорення інференсу моделей, оптимізованих під фреймворк Intel OpenVINO, що дозволяє значно знизити затримки та енергоспоживання в порівнянні з класичними CPU- та GPU-обчисленнями. Завдяки компактному розміру і низькому тепловому пакету, модуль ідеально підходить для використання в системах відеоспостереження, IoT-пристроях, робототехніці та інших вбудованих

рішеннях, де обмежені апаратні ресурси і необхідна висока швидкість обробки зображень[5].

1.2 Існуючі системи розпізнавання

Рішення задачі виявлення об'єктів на зображенні та відеофайлі використовується в багатьох областях діяльності, таких як контроль дорожнього руху, охоронні системи, промисловість, розпізнавання текстових документів, кінематограф, робототехніка, медичні дослідження, аналіз космічних знімків і багатьох інших. Задачі комп'ютерного зору є актуальними, оскільки перераховані сфери в даний час важливі і активно розвиваються.

У кінематографі, анімації, будівництві та дизайні інтер'єру комп'ютерний зір застосовується для генерації 3D-моделей об'єктів.

У промисловості технології Computer Vision потрібні для забезпечення безпеки робітників. На багатьох підприємствах проводять ідентифікацію особистості. На небезпечних виробництвах системи контролюють носіння робітниками засобів індивідуального захисту – спецодягу, касок, рукавичок та ін.

При розробці ШІ-рішень в області медицини велику роль мають технології машинного зору. Не так давно компанія Philips розробила нейромережний алгоритм, який дозволяє покращувати якість рентгенівських зображень і зображень комп'ютерної томографії. Він прибирає шуми, спотворення, дисторсію, що дозволяє обстежуваним набагато менше часу перебувати в томографі і отримувати менше променевого навантаження [5].

Зараз, у зв'язку зі сформованою епідеміологічною ситуацією, в багатьох установах є системи комп'ютерного зору, які контролюють дотримання людьми маскового режиму: вони аналізують відеофайли і працюють в режимі реального часу, виявляють порушення, пов'язані з відсутністю масок або інших засобів захисту органів дихання.

Системи розпізнавання присутні навіть в кожному сучасному

смартфоні. У сучасних смартфонах Google Pixel, зокрема моделей Pixel 8 (2023) та Pixel 9 (2024), реалізовано передові алгоритми комп'ютерного зору, що дозволяють користувачам обробляти зображення безпосередньо під час зйомки. Система автоматично розпізнає обличчя, оцінює освітлення, кольоровий баланс, контраст і динамічний діапазон. Технологія HDR+ забезпечує високоякісні результати навіть за складних умов. Також реалізовано функції на основі штучного інтелекту: «Best Take» для вибору найкращого кадру, «Zoom Enhance» для відновлення чіткості під час масштабування, «Night Sight» для нічної зйомки та Pro Mode з ручним налаштуванням параметрів. Ці інновації дозволяють значно покращити якість фотографій і відео навіть без професійного обладнання[7].

В області контролю дорожнього руху теж використовуються технології Computer Vision. В Україні розроблено декілька апаратно-програмних комплексів для розпізнавання автомобільних державних номерних знаків (ДНЗ). Найпоширенішими є «Kobi Software», «INNI Tech» і «vCloud.ai» [8-10].

Kobi Software (ZetPro VMS) – система від української компанії Kobi Software призначена для розпізнавання державних номерних знаків транспортних засобів у режимі реального часу. Вона працює як частина відеоменеджмент-системи ZetPro VMS, має високу точність обробки відео- і фотопотоків, підтримує роботу в умовах низької освітленості та дозволяє інтеграцію з іншими системами безпеки. Призначена для використання на паркінгах, платних дорогах, АЗС та об'єктах критичної інфраструктури, де важливо відстежувати в'їзд і виїзд транспортних засобів.

INNI Tech – українська розробка, яка використовує технології штучного інтелекту та згорткових нейронних мереж для точного розпізнавання номерних знаків. Платформа дозволяє інтегрувати власні моделі машинного навчання, що робить її придатною для експериментів та академічних досліджень. INNI Tech підтримує обробку відео та фото в реальному часі, дозволяючи реалізувати проекти на базі камер спостереження або мобільних пристроїв. Система широко використовується в міському моніторингу,

контролі трафіку та автоматизованих системах паркування.

vCloud.ai – хмарна система відеоаналітики, що включає модуль розпізнавання автомобільних номерів з використанням штучного інтелекту. Вона дозволяє аналізувати відеопотоки з великої кількості камер, виявляти та фіксувати номерні знаки транспортних засобів, зберігати інформацію та передавати її до серверів чи баз даних. Система орієнтована на муніципальні та комерційні структури, забезпечуючи автоматизацію процесів контролю доступу, безпеки та відеоспостереження в режимі 24/7

Чому підходить: Хмарне рішення, яке використовує нейронні мережі для аналізу відео. Добре ілюструє сучасний підхід до віддаленої обробки зображень

1.3 Проблеми систем розпізнавання

До основних проблем систем розпізнавання об'єктів на зображеннях і відеофайлах відносять наступні:

- низька якість зображення (причинами можуть бути зламана камера, невірні установки знімальної апаратури, велика відстань від камери до об'єкта, неправильна обробка фото та ін.),
- погане освітлення, наявність тіней,
- розмитість зображення (ця проблема найчастіше присутня в системах, які проводять відеофіксацію рухомих об'єктів),
- наявність предмета, який повністю або частково закриває собою досліджуваний об'єкт.

Крім перерахованих вище проблем, властивих всім типам систем розпізнавання, можна виділити додатковий ряд проблем, які зустрічаються в системах розпізнавання ДНЗ:

- відмінності між номерними знаками різних країн. Цю проблему можна вирішити розширенням бази даних шляхом додавання в неї прикладів номерних пластин інших країн,

- наявність причепа. В цьому випадку необхідно розпізнати дві номерні пластини – основного транспортного засобу і його причепа,
- наявність буксира (аналогічно проблемі вище),
- зміна водієм автомобіля смуги руху в момент зчитування камерою його ДНЗ. Ця проблема властива системам, що працюють в режимі реального часу, – наприклад, фото- і відеокамерам, що фіксують порушення правил дорожнього руху на дорогах,
- закриття деяких символів номерного знака брудом, багажем, буксирними гаками та іншими предметами.

Чим з більшою кількістю перерахованих вище проблем справляється система, тим якіснішою вона є.

2 МЕТОДИ РОЗПІЗНАВАННЯ ГРАФІЧНИХ ОБРАЗІВ

2.1 Основні методи розпізнавання образів

Образом, як правило, називають відтворення деякого об'єкта або групи елементів, об'єднаних між собою за будь-якою ознакою і мають характерні властивості. Досліджуваний об'єкт відносять до того чи іншого образу за встановленими заздалегідь правилами. Кожен об'єкт може бути розглянутий як багатовимірний вектор, координати якого відображають значення його характеристик.

Подібності і відмінності між об'єктами можна визначити за допомогою поняття відстані між векторами - метрики. Вектори є набори характеристик об'єктів. Чим менша відстань між векторами, тим більше схожими є відповідні їм об'єкти. Відстань обчислюється за заздалегідь заданим способом, який залежить від структури об'єктів.

У теорії розпізнавання образів виділяють три основні групи методів розпізнавання:

- методи перебору,
- методи з використанням характеристик об'єктів,
- методи, що використовують нейронні мережі.

У методах перебору досліджуваний об'єкт порівнюється з образами, що зберігаються в базі даних.

Методи з використанням характеристик об'єктів працюють складніше. При їх застосуванні здійснюють аналіз образу. Наприклад, при виявленні об'єктів на зображенні вирішують завдання визначення області, яка зайнята об'єктом, завдання виділення країв, виконують сегментацію зображення, визначають геометричні розміри об'єкта і інші його характеристики.

Методи, що використовують штучні нейронні мережі, вважаються одними з найскладніших. Для кожної окремої задачі необхідно підбирати

спеціальну структуру нейронної мережі. Структура залежить від особливостей задачі і від виду необхідного рішення. Однак перевага таких методів полягає в їх високій продуктивності та ефективності.

2.2 Розпізнавання за ознаками

Розпізнавання образів за набором їх ознак відноситься до групи методів з використанням характеристик об'єктів. Об'єкти описуються набором ознак (характеристик). Інакше кажучи, об'єкт представляється вектором x , що належить n -вимірному простору, і компоненти цього вектора є характеристиками об'єкта.

При розпізнаванні образів на зображеннях вектори можуть включати в себе:

- геометричні розміри об'єкта: периметр, площа, висота, ширина, довжина,
- структуру: контур і його розміри, положення країв,
- інші характеристики [3].

Компоненти векторів мають числові значення. Класифікатор за допомогою такого вектора відносить об'єкт x до того чи іншого класу відповідно до розбиття n -вимірного простору. При цьому він використовує метрику.

Найчастіше застосовують евклідову метрику, обчислювану за формулою (2.1):

$$d_E(x_i x_j) = \sqrt{\sum_k (x_{ik} x_{jk})^2}. \quad (2.1)$$

де x_i , x_j – вектори; k – індекс компоненти вектора. Ще часто використовують «манхеттенську відстань» (відстань міських кварталів) – метрику, введenu Германом Мінковським. Вона дорівнює сумі модулів різниць компонент векторів (формула (2.2)):

$$d_M(x_i, x_j) = \sum_k |x_{ik} - x_{jk}|. \quad (2.2)$$

Також вводять в розгляд відстань Геммінга, що дорівнює числу позицій, в яких вектори розрізняються. Якщо вектори n -вимірні, то очевидно, що відстань Геммінга не перевищуватиме n (формула (2.3))

$$d_H(x_i, x_j) \leq n. \quad (2.3)$$

Запит виконано. Вектори називаються сусідніми, якщо відстань Геммінга між ними дорівнює одиниці.

2.3 Структурні методи розпізнавання

У методі розпізнавання образів за набором їх ознак, описаному вище, використовуються тільки числові значення. Для об'єктів зі складною структурою цього може бути недостатньо. Тоді необхідно до наявних характеристик об'єкта додати додаткові. Такі методи і називаються структурними, оскільки структура об'єкта не може бути описана тільки за допомогою числових параметрів. Об'єкт, як правило, розбивають на простіші елементи і визначають відносини між ними.

Складні образи складаються з більш простих підобразів. Опис відношень між підобразами можна представити у вигляді ієрархії, наприклад, в графовій структурі. Вершинами графа є частини і елементи образу, а ребрами - відносини між ними. Відносини можуть описуватися логічними або математичними функціями.

Можна сказати, що елементи, відносини і орієнтований граф в сукупності утворюють один з типів семантичних мереж [4] - функціональну мережу. Формат функціональних мереж передбачає використання цілих

образів, які пов'язані між собою. Зв'язки представлені у вигляді графа. У графах допускається множинність зв'язків (ребер) між парами вершин.

Структурний метод розпізнавання дає можливість описувати велику кількість складних об'єктів, використовуючи відносно невелику базу з нескладними образами. Розпізнавання проводиться в три етапи. На першому етапі здійснюється підготовка об'єкта до розпізнавання – перевірка його наявності на зображенні, визначення меж і країв. Далі, на другому етапі, об'єкт представляють у вигляді структури: проводиться сегментація, об'єкт ділиться на частини – елементи, між ними визначаються відносини. Елементи і відносини беруться з бази правил, що відображає семантику предметної області. На третьому етапі виконується аналіз отриманої структури. На цьому кроці приймається рішення, чи є представлення об'єкта вірним, і в залежності від результату рішення об'єкт або виключається з розгляду, або визначається в клас образів.

2.4 Статистичні методи

Методи, засновані на теоремі Байєса, класифікують об'єкти, враховуючи значення ймовірності належності цих об'єктів до різних класів. В якості гіпотез виступають гіпотези належності об'єкта до певного класу. Теорема Байєса представима у вигляді формули (2.4):

$$P(H_i|D) = \frac{P(D|H_i)P(H_i)}{P(D)}. \quad (2.4)$$

У формулі (2.4) ймовірність $P(H_i|D)$ називається апостеріорною. Такі ймовірності обчислюються для кожного класу. Апостеріорна ймовірність є умовною і залежить від даних, отриманих з досвіду.

Ймовірність $P(D|H_i)$ називається правдоподібністю отримання вже відомих даних у разі, якщо розглянута гіпотеза вірна.

$P(H_i)$ називається апіорним переконанням і виражає ймовірність гіпотези до введення в розгляд дослідних даних.

$P(D)$ – це ймовірність даних, які спостерігаються незалежно від гіпотези [5]. Іноді останню ймовірність не використовують, оскільки обчислювати її скрутно. Замість цього порівнюють максимальні величини чисельника правої частини формули (2.4).

Статистичні методи застосовні у випадку, коли всі класи володіють власними значеннями ознак належним їм об'єктів. Якщо ж класи перетинаються в деяких значеннях ознак, то можливі помилки розпізнавання.

2.5 Штучні нейронні мережі

2.5.1 Концептуальна модель

Одним з найбільш часто використовуваних класів алгоритмів для вирішення задач машинного зору є нейронні мережі.

Штучна нейронна мережа (ШНМ, НМ) – це паралельно розподілена структура обробки інформації, яка складається з нейронів, пов'язаних між собою. Модель нейронної мережі в програмуванні є машинною інтерпретацією головного мозку. Центральна нервова система людини представлена спинним і головним мозком, функціональність яких здійснюється за допомогою величезної кількості нейронів, які пов'язані між собою синаптичними зв'язками і можуть передавати інформацію за допомогою імпульсів.

Перцептрон – це математична модель сприйняття інформації людським мозком, запропонована Ф. Розенблаттом в середині минулого століття [6]. Перцептрон став однією з перших моделей нейронних мереж, яку згодом стали активно розвивати і покращувати.

НМ володіють здатністю до самонавчання. Інакше кажучи, вони можуть видавати результат на основі отриманого досвіду, узагальнюючи наявні прецеденти на нові випадки. ШНМ складаються з вхідного, прихованих і вихідного шарів (рисунок 2.1). Кожен шар містить певну кількість нейронів.

Шари нейронної мережі

Процес навчання нейронної мережі є ітеративним, його кроки називають епохами. Для навчання НМ найчастіше використовують методи навчання з учителем, наприклад, метод зворотного поширення помилки і його модифікації. Нерідко використовують і навчання без вчителя. Тип навчання головним чином залежить від вирішуваної задачі. Навчання з учителем відбувається при наявності повного набору розмічених даних. У навчальному наборі кожному прикладу відповідає рішення, яке мережа повинна отримати. Різниця між правильним рішенням і отриманим є помилкою, яку необхідно усувати за допомогою нової настройки параметрів. Навчання без учителя проводиться без контролю розробника над процесом. На вхід подається набір даних, і нейронна мережа намагається самостійно знайти взаємозв'язки. При навчанні без учителя у розробника відсутні правильні рішення, які мережа повинна отримати. Одним з достоїнств НМ є вирішення задач в умовах невизначеності. Здатність до самонавчання дозволяє мережам шукати вирішення задач з невідомими закономірностями і залежностями між вхідними і вихідними даними. Згортова штучна нейронна мережа (ЗШНМ, ЗНМ) - це архітектура штучних нейронних мереж, яка входить до складу технологій глибокого навчання. Такі мережі націлені на розпізнавання образів. Їх модель в програмуванні заснована на особливостях роботи клітин зорової кори головного мозку. Для вирішення задач комп'ютерного зору широко використовується технологія згорткових нейронних мереж. Перевагою такого підходу є те, що такі мережі є одним з кращих алгоритмів з розпізнавання та класифікації зображень, стійкі до повороту і зсувів зображень, мають набагато менше настроюваних ваг в порівнянні зі звичайною нейронною мережею. Ще однією перевагою є можливість реалізації паралельно працюючих алгоритмів [3], що дозволяє підвищити продуктивність системи і збільшити швидкість

обчислень. Згорткові мережі працюють за тим же принципом, що і звичайні нейронні мережі, тільки крім множення матриць в них присутня операція згортки. Згортка - це лінійна операція, що застосовується до двох функцій дійсного аргументу і повертає третю, яка характеризує подібність однієї функції з відбитою і зрушеною копією іншої функції [7]. Ще одна відмінність згорткових мереж - це розріджена зв'язність [8]. У звичайному багатошаровому перцептроні кожен нейрон наступного шару пов'язаний з усіма нейронами попереднього шару, і всі зв'язки мають свої ваги в матриці ваг. А в ЗНМ використовується невелика матриця ваг, яка на кожному кроці рухається по оброблюваному шару. Матрицю ваг називають ядром згортки, в мережі ядер кілька. Ядра кодують наявність будь-яких ознак на зображенні - горизонтальних і вертикальних ліній, ліній під кутом, дуг, складних фігур (еліпсів, трикутників, квадратів та інших). Наступний шар, що формується в результаті згортки однієї з матриць ваг, буде відображати існування певної ознаки в обробленому шарі. Таким чином формується карта ознак, що представляє собою масив матриць. Кожна карта містить в собі синаптичне ядро або фільтр. Фільтр показує наявність певної ознаки. Прохід кожним набором ваг становить свій екземпляр карти ознак, роблячи нейронну мережу багатоканальною (тобто багато незалежних карт ознак на одному шарі) [9]. При виконанні операції згортки вікно з розмірністю ядра проходить із заданим кроком всю область зображення, на кожному кроці проходження множить вміст вікна на ядро, результат підсумовує і записує в матрицю результату - чергову карту ознак. Розміри карт ознак одного згорткового шару однакові. Операція підвибірки або пулінг (англ. Pooling) необхідна для зменшення масштабу карт ознак. Найчастіше використовується операція MaxPooling - відбір найбільших значень: з декількох сусідніх нейронів карти вибирають максимальний і приймають його за один нейрон нової карти ознак меншої розмірності. Завдяки цьому можна значно знизити обсяг використовуваної пам'яті і прискорити процес подальших обчислень (рисунки 2.1).

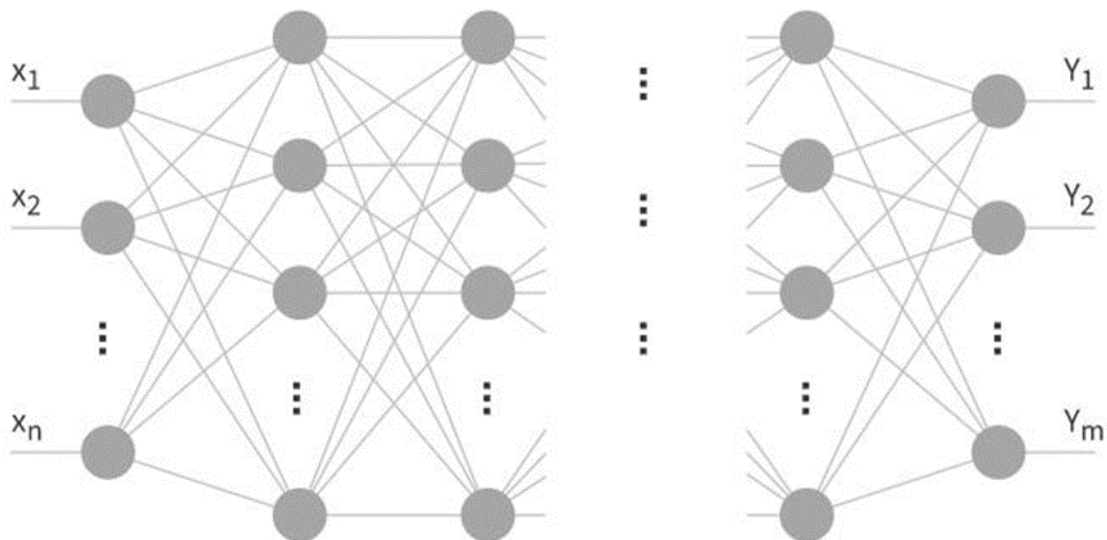


Рисунок 2.1 – Шари нейронної мережі

Процес навчання нейронної мережі є ітеративним, його кроки називаються епохами.

Для навчання НМ (нейронних мереж) найчастіше використовують методи навчання з учителем, наприклад, метод зворотного поширення помилки та його модифікації. Нерідко використовують і навчання без учителя. Тип навчання головним чином залежить від вирішуваної задачі.

Навчання з учителем відбувається за наявності повного набору розмічених даних. У навчальному наборі кожному прикладу відповідає рішення, яке мережа повинна отримати. Різниця між правильним рішенням та отриманим являє собою помилку, яку необхідно усувати за допомогою нової настройки параметрів.

Навчання без учителя проводиться без контролю розробника над процесом. На вхід подається набір даних, і нейронна мережа намагається самостійно знайти взаємозв'язки. При навчанні без учителя у розробника відсутні правильні рішення, які мережа повинна отримати.

Однією з переваг НМ є вирішення задач в умовах невизначеності. Здатність до самонавчання дозволяє мережам шукати рішення задач з невідомими закономірностями та залежностями між вхідними та вихідними даними.

Згорткова штучна нейронна мережа (ЗШНМ, ЗНМ) – це архітектура штучних нейронних мереж, яка входить до складу технологій глибокого навчання. Такі мережі націлені на розпізнавання образів. Їхня модель у програмуванні ґрунтується на особливостях роботи клітин зорової кори головного мозку.

Для вирішення задач комп'ютерного зору широко використовується технологія згорткових нейронних мереж. Перевагою такого підходу є те, що такі мережі є одним з найкращих алгоритмів з розпізнавання та класифікації зображень, стійкі до повороту та зсувів зображень, мають значно менше налаштовуваних ваг порівняно зі звичайною нейронною мережею. Ще однією перевагою є можливість реалізації паралельно працюючих алгоритмів [3], що дозволяє підвищити продуктивність системи та збільшити швидкість обчислень.

Згорткові мережі працюють за тим же принципом, що й звичайні нейронні мережі, тільки крім множення матриць у них присутня операція згортки. Згортка – це лінійна операція, що застосовується до двох функцій дійсного аргументу та повертає третю, яка характеризує подібність однієї функції з відбитою та зсунутою копією іншої функції [9]. Ще одна відмінність згорткових мереж – це розріджена зв'язність [10].

У звичайному багатосаровому перцептроні кожен нейрон наступного шару пов'язаний з усіма нейронами попереднього шару, і всі зв'язки мають свої ваги в матриці ваг. А в ЗНМ використовується невелика матриця ваг, яка на кожному кроці рухається по оброблюваному шару.

Матрицю ваг називають ядром згортки, у мережі ядер декілька. Ядра кодують наявність будь-яких ознак на зображенні – горизонтальних і вертикальних ліній, ліній під кутом, дуг, складних фігур (еліпсів, трикутників, квадратів та інших). Наступний шар, що формується в результаті згортки однієї з матриць ваг, відобразить існування певної ознаки в обробленому шарі. Таким чином формується карта ознак, що являє собою масив матриць.

Кожна карта містить у собі синаптичне ядро або фільтр. Фільтр показує

наявність певної ознаки. Прохід кожним набором ваг становить свій екземпляр карти ознак, роблячи нейронну мережу багатоканальною (тобто багато незалежних карт ознак на одному шарі) [10].

При виконанні операції згортки вікно з розмірністю ядра проходить із заданим кроком всю область зображення, на кожному кроці проходу множить вміст вікна на ядро, результат підсумовує та записує в матрицю результату – чергову карту ознак. Розміри карт ознак одного згорткового шару однакові.

Операція підвибірки або пулінг (англ. Pooling) необхідна для зменшення масштабу карт ознак. Найчастіше використовується операція MaxPooling – відбір найбільших значень: з декількох сусідніх нейронів карти вибирають максимальний і приймають його за один нейрон нової карти ознак меншої розмірності. Завдяки цьому можна значно знизити обсяг використовуваної пам'яті та прискорити процес подальших обчислень.

Описану вище процедуру в графічному представленні можна побачити на рисунку 2.1.

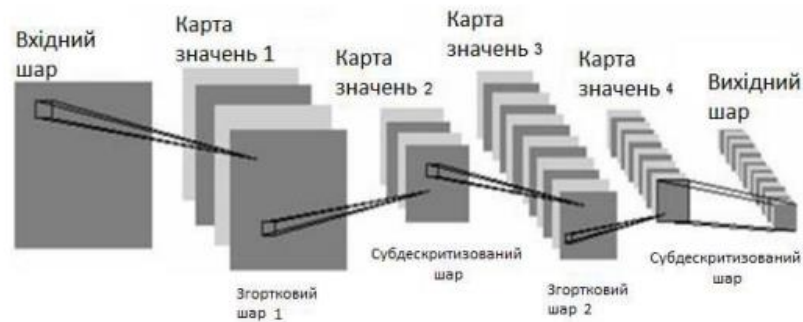


Рисунок 2.1 – Схема роботи згорткової мережі

Сигнал проходить через певну кількість шарів (на рисунку: один вхідний шар, що представляє собою безпосередньо саме вхідне зображення, чотири приховані шари, один вихідний шар), де чергуються згортка і підвибірка (пулінг). При цьому складаються карти ознак. Їх кількість збільшується з кожним шаром, але одночасно зменшується роздільна здатність карт [11].

На виході виходить великий набір каналів, що містять невелику

кількість даних, які відображають ознаки, виявлені на вхідному зображенні. Потім, як правило, ці дані передаються на звичайну повносвязную нейронну мережу, яка формує кінцевий вихідний сигнал. У підсумку вихідне зображення проходить через множинну фільтрацію.

Ядра згортки формуються за допомогою навчання нейронної мережі методом зворотного поширення помилки. Процеси згортки, що виконуються по кожній карті ознак, відбуваються паралельно, що є істотною перевагою з погляду часу роботи мережі.

Кількість карт визначається вимогами до відповіді вирішуваної задачі. Якщо буде велике число карт, то покращиться якість розпізнавання, але в той же час підвищиться складність обчислень. Найчастіше береться співвідношення 1:2 - на одну карту попереднього шару припадає дві карти нового шару.

Недоліком ЗНМ є наявність великої кількості змінюваних параметрів - кількість шарів, кількість і розмірність ядер згортки, крок зсуву ядра при обробці шару, параметри підвибірок, а також параметри вихідної повносвязной мережі.

2.5.2 Функції активації нейронів

У нейромережах функції активації (ФА) визначають вихідний сигнал нейрона в залежності від вхідного сигналу або набору вхідних сигналів.

Виділяють три основні типи функцій активації:

- порогові функції або функції одиничного стрибка,
- кусково-лінійні функції,
- сигмоїдальні функції.

Порогова функція в простій формі є двійковою: нейрон або збуджений, або ні. Така функція є східчасту функцію Хевісайда і обчислюється за формулою (2.5). Вона підходить для бінарної класифікації.

$$\varphi(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (2.5)$$

Кусково-лінійна функція в загальному випадку представляється у вигляді формули (2.6):

$$\varphi(x) = \begin{cases} 1, & x \geq 0,5 \\ |x|, & -0,5 < x < 0,5 \\ 0, & x \leq -0,5 \end{cases} \quad (2.6)$$

Але частіше використовують функцію ReLU, так звану випрямлену лінійну одиницю (формула (2.7)):

$$f(x) = \max(0, x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (2.7)$$

Функція ReLU має кілька переваг: швидке обчислення похідної, легка оптимізація. Але також у неї є недолік, який називають проблемою вмираючого ReLU: якщо нейрон став від'ємним, то він не відновиться і, відповідно, навчатися не буде. У такому випадку зазвичай збільшують кількість нейронів. Функція ReLU та її модифікації є найбільш широко використовуваними в згорткових нейронних мережах.

Сигмоїдальна функція – це неперервна і диференційована функція, яка підтримує баланс між лінійною і нелінійною поведінкою [11]. На вході вона приймає дійсне число від мінус нескінченності до плюс нескінченності, на виході видає дійсне число в діапазоні від 0 до 1. Сигмоїда має вигляд формули (2.8). Незважаючи на корисні властивості неперервності і диференційованості, вона має недолік, що полягає в тому, що при насиченні функції з боку нуля або одиниці градієнт буде близький до нуля.

$$\sigma(x) = \frac{1}{1+e^{-x}}. \quad (2.8)$$

Функція Softmax (зважена сигмоїда) використовується для множинної класифікації. Така функція нормалізує N-мірний вектор довільних значень в N-мірний вектор, в якому кожна координата приймає значення в діапазоні від 0 до 1. Як правило, Softmax застосовують в останньому шарі нейронної мережі. Координати $j = \overline{1, N}$ обчислюються за формулою (2.9). Координата j отриманого вектора відображає ймовірність приналежності об'єкта до класу з номером j .

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{n=1}^N e^{z_n}}. \quad (2.9)$$

Функція Softmax (зважена сигмоїда) використовується для множинної класифікації.

2.5.3 LeNet

LeNet – це перша структура згорткової нейронної мережі, розроблена Яном ЛеКуном наприкінці минулого століття. Тоді, за вкрай обмежених ресурсів, вона дозволила побудувати і успішно навчити НМ, що розпізнає і класифікує чорно-білі обриси цифр. Це був прорив у розпізнаванні цифр і букв.

Мережа складається з 7 шарів. На вході – чорно-біле зображення цифри розміром 32 на 32 пікселі. Згортковий шар складається з 6 фільтрів 5×5 , вони виділяють на зображенні прості форми (ламані, смуги та ін.). Після застосування фільтрів виходить 6 проекцій 28×28 (крок згортки дорівнює 1, формула обчислення розміру вихідного зображення: $32 \text{ мінус } 5 \text{ плюс } 1 \text{ дорівнює } 28$), тобто розмір вихідного зображення зменшується. Далі застосовується підвибірка MaxPooling 2×2 . Розмірність зображень знову зменшується і стає 14×14 . Далі йде другий шар згортки – 16 ядер 5×5 . Шар

підвибірці аналогічний попередньому. Розмірність стає 10×10 .

Наступні два шари – звичайні шари: 120 нейронів з функцією активації $\tanh()$ і 84 нейрони з цією ж функцією. Вихідний (останній) шар має 10 нейронів, оскільки класифікує 10 класів цифр.

На обладнанні минулого століття мережа навчалася кілька годин, навчання пройшло успішно. Вважається, що LeNet поклала початок згортковим НМ.

LeNet та її модифікації – здатні моделі, їх легко навчати, вони мають високу продуктивність. Але ці НМ не можна застосовувати до зображень з високою роздільною здатністю.

У 2012 році була опублікована архітектура ЗНМ AlexNet, розроблена Алексом Кріжевським у співпраці з Іллею Суцкевером і Джеффри Хінтоном [11]. Її схема зображена на рисунку 2.2.

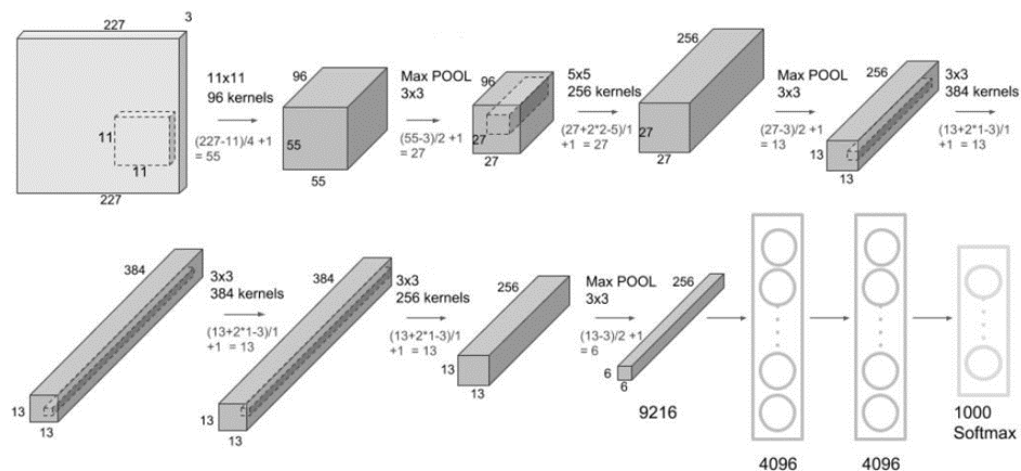


Рисунок 2.2 – Мережа AlexNet

Мережа дозволяє працювати з великими повноколірними зображеннями.

AlexNet з великим відривом виграла конкурс з розпізнавання зображень ImageNet Large Scale Visual Recognition Challenge з кількістю помилок 15,3 відсотка проти 26,1 відсотка в іншого рішення, що посіло друге місце.

Архітектура містить 8 шарів. Перші 5 є згортковими, в них як функція

активації використовується ReLU. За рахунок використання цієї ФА швидкість роботи збільшилася в 6 разів. Архітектуру AlexNet використовують досі, наприклад, вона реалізована в Python-бібліотеках глибокого навчання TensorFlow і Keras.

Вихідне зображення має роздільну здатність 227×227 пікселів. Є 3 кольорних канали: червоний, зелений, блакитний (RGB), на кожен канал накладається фільтр розміру 11×11 . Розмірність в результаті роботи шарів змінюється наступним чином: 227×227 , 55×55 , 27×27 , 13×13 , 6×6 .

Після 5 згорткових шарів йдуть 2 пов'язаних шару, що містять по 4096 нейронів. Вихідний шар, що використовує ФА Softmax, здатний класифікувати об'єкти на 1000 класів.

Архітектура AlexNet послужила основою для багатьох сучасних згорткових мереж. Концепцію архітектури стали активно розвивати.

У 2014 році як поліпшення архітектури AlexNet була запропонована мережа VGG. Її розробила група комп'ютерного зору Оксфордського університету (Visual Geometry Group) разом з дослідниками з Google DeepMind [12].

У цій мережі великі фільтри замінені на деяку кількість маленьких. Замість фільтрів 11×11 і 5×5 в перших двох згорткових шарах кілька разів використовуються фільтри розміру 3×3 . Після каскаду згорток використовується та ж конфігурація пов'язаних шарів, що і в мережі AlexNet.

Завдяки введеному нововведенню вдалося зменшити кількість ваг в мережі. VGG істотно перевершила попередні покоління моделей в продуктивності.

ResNet - це скорочення від Residual Network (залишкова мережа) [13]. Така мережа дозволяє вирішити проблему погіршення точності передбачення зі збільшенням глибини структури мережі.

Проблема зникаючого градієнта полягає в тому, що на якомусь етапі навчання мережі він стає дуже малим, що не дозволяє змінюватися вагам. У

залишкових неймережах є ярлики з'єднання, в основі роботи яких лежить відображення ідентичності. Мережа дозволяє на деяких етапах навчання пропускати шари. Також у ResNet немає пов'язаних шарів в кінці, присутній тільки один шар, що передбачає вихідні класи. Навчання триває приблизно 60 000 епох.

В результаті експериментів з ResNet було встановлено, що неймережі з великою глибиною можна навчати без погіршення точності.

3. ПРОГРАМНА РЕАЛІЗАЦІЯ АЛГОРИТМУ

3.1 Постановка задачі

Практичною задачею роботи була розробка алгоритму розпізнавання символів на державних номерних знаках, що використовує технологію згорткових нейронних мереж. Для створення і навчання нейронної мережі було потрібно визначитися з інструментами розробки, вибрати оптимальну структуру мережі, налаштувати її параметри. При вирішенні задачі була використана структура ЗНМ LeNet. Для навчання мережі був знайдений набір даних з різними зображеннями 21 символу — 10 цифр і 11 букв. Нейронна мережа навчилася на цьому наборі даних розпізнавати окремі символи на ДНЗ. Ідея алгоритму полягала в наступному. Алгоритм отримує на вхід зображення номерного знака транспортного засобу, потім розбиває зображення на сегменти - тобто, кожен символ номерного знака стає окремим зображенням. Далі відбувається робота нейронної мережі, яка розпізнає кожен виділений символ окремо, після чого алгоритм збирає в один рядок розпізнані символи номера транспортного засобу і видає цей рядок як відповідь на поставлену задачу.

1.2 Засоби розробки рішення Основним інструментом розробки був обраний мову програмування Python. Він є високорівневою, об'єктно-орієнтованою мовою, що включає можливість інтегруватися з іншими мовами програмування. Його інтерпретатор має відкритий вихідний код. В даний час мова широко використовується у сфері аналізу даних, машинному навчанні, веб-розробці та інших областях. В якості середовища розробки було вибрано інтерактивне середовище Jupyter Notebook, в якому можна відразу бачити результати виконання коду і його окремих частин. У Jupyter код можна розділяти на фрагменти і виконувати їх в довільному порядку. Ще одна перевага — код програми може об'єднуватися з документацією та зображеннями. Версія 3 мови Python надає велику кількість бібліотек для

фахівців Data Science. Найбільш популярними є наступні: - numpy. Вона надає реалізації математичних операцій і алгоритмів, скомпільованих на мові програмування C. Ядром бібліотеки є багатомірний масив даних, з яким і працює програміст [12]; - matplotlib. Ця бібліотека є засобом візуалізації даних. З її допомогою можна формувати графіки і діаграми, двомірну і тривимірну графіку, статичні і анімовані зображення [13]; - PIL (Python Imaging Library). Ця бібліотека створена для роботи з растровою графікою, вона підтримує багато поширених формати: BMP, JPEG, PNG, GIF та інші. Корисною можливістю бібліотеки є перетворення зображень - масштабування, застосування фільтрів, малювання і т. д.; - keras. Вона є популярним інструментом машинного навчання. Бібліотека поставляється з різними наборами даних, які можна знайти в репозиторії GitHub; - opencv (Open-Source Computer Vision Library). Ця бібліотека зберігає алгоритми обробки зображень.

3.2 Засоби розробки рішення

Основним інструментом реалізації проєкту було обрано мову програмування Python. Це високорівнева об'єктно-орієнтована мова, яка підтримує інтеграцію з іншими мовами та має відкритий вихідний код. Python отримав широке поширення в галузях обробки даних, машинного навчання, наукових обчислень і веброботи.

Для написання коду використовувалося інтерактивне середовище Jupyter Notebook, що дозволяє запускати окремі частини коду та одразу переглядати результати їх виконання. Середовище зручне тим, що дозволяє поєднувати фрагменти коду з текстовими поясненнями, графіками, таблицями та зображеннями, що особливо важливо під час дослідницької роботи.

У розробці було використано Python 3, який надає доступ до широкого спектру бібліотек для аналізу даних і побудови моделей машинного навчання. Серед найважливіших:

- NumPy – бібліотека для високопродуктивних математичних операцій, що базується на багатовимірних масивах та реалізована на мові C;
- Matplotlib – засіб для побудови графіків та візуалізації статистичних даних;
- PIL (Python Imaging Library) – бібліотека для роботи з растровими зображеннями, яка підтримує формати JPEG, PNG, BMP, GIF тощо, і забезпечує можливості масштабування, обробки фільтрами та малювання;
- Keras – популярна високорівнева бібліотека для створення і навчання моделей штучних нейронних мереж, сумісна з TensorFlow і має вбудовані набори даних;
- OpenCV (Open Source Computer Vision Library) – потужна платформа для комп'ютерного зору, яка включає алгоритми обробки, фільтрації, сегментації зображень та багато іншого.

Цей набір інструментів забезпечує ефективну розробку, тренування та тестування моделей глибокого навчання, зокрема згорткових нейронних мереж, у рамках завдання з розпізнавання номерних знаків.

3.3 Набір даних для навчання НМ

Для реалізації задачі розпізнавання символів на зображеннях автомобільних номерів було використано набір даних Train_Cells. Джерелом доступу до нього став популярний майданчик для обміну наборами даних – Kaggle, який належить компанії Google.

Цей датасет містить окремі зображення цифр та літер, які було попередньо вирізано з фотографій державних реєстраційних номерів транспортних засобів (формату України). Усі зображення згруповано за класами (відповідними символами) в окремі папки та збережено у форматі BMP.

Зазначений набір даних виконував функцію тренувальної вибірки, тобто слугував основою для навчання нейронної мережі. Кожне зображення у

колекції має заздалегідь відомий клас (літеру або цифру), що дозволяє застосовувати методи навчання з учителем. Частина прикладів з цього набору наведено на рисунку 3.1.



Рисунок 3.1 – Частина навчального набору

Вид реєстраційних номерних знаків (НЗ) України — це класифікація державних номерних знаків транспортних засобів за формою, кольором, призначенням і категорією власника згідно з технічними стандартами та нормативами, прийнятими в Україні [14].

- формат номерного знаку:
- стандартний зразок (після 2004 року): AA 0000 AA;
- перші дві літери – код регіону реєстрації (напр., АХ – Харків);
- далі – чотири цифри;
- останні дві літери – серія.

Формат має вигляд дві літери, чотири цифри, 2 літери (рисунок 3.2).

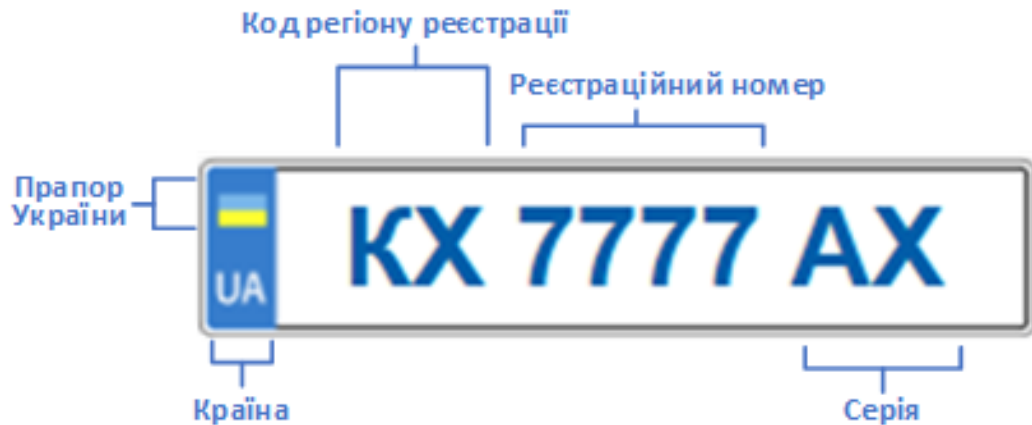


Рисунок 3.2 – Формат номерного знаку

Варто підкреслити, що навчальний датасет складався не з повноцінних зображень номерних знаків, а з окремих символів (цифр і літер), вирізаних із реальних номерних пластин. Кожне зображення представляє собою ізольований символ, що належить до певного класу.

У процесі роботи згортова нейронна мережа не аналізує номерний знак цілком, а проводить класифікацію кожного символу окремо. Далі — на етапі постобробки – система об'єднує послідовність розпізнаних символів у повний реєстраційний номер транспортного засобу, відтворюючи його у відповідному форматі (наприклад, KX 7777 AX).

Такий підхід значно спрощує структуру навчального процесу та підвищує точність класифікації, оскільки кожен фрагмент аналізується незалежно та з меншим впливом на загальне спотворення або зашумлення зображення.

3.4 Структура алгоритму

Алгоритм розпізнавання державного номерного знака складається з наступних етапів:

а) навчання нейронної мережі. На першому етапі здійснюється тренування згортової нейронної мережі (ЗНМ) на наборі даних Train_Cells з

метою навчання розпізнавання окремих символів (літер та цифр) на зображенні;

б) отримання знімка НЗ. Алгоритм отримує вхідне зображення державного номерного знака транспортного засобу;

в) сегментація НЗ на окремі символи. За допомогою інструментів бібліотеки OpenCV виконується розділення зображення номерного знака на окремі символи. Алгоритм виявляє контури кожної літери та цифри, після чого кожен символ зберігається у вигляді окремого зображення. Детальніше метод сегментації описано нижче;

г) розпізнавання символів нейронною мережею. Отримані на попередньому етапі зображення символів послідовно подаються на вхід згорткової нейронної мережі. Мережа здійснює класифікацію кожного символу, а результати розпізнавання об'єднуються у єдиний рядок, який видається як фінальний результат.

3.5 Сегментація зображень

Для сегментації знімка номерного знака було використано інструменти бібліотеки OpenCV.

Метод Thresholding (Порогова обробка)

Thresholding – це метод сегментації, призначений для перетворення зображення у бінарне представлення [14]. Він поділяється на два типи:

- простий поріг фіксоване значення для всього зображення;
- адаптивний поріг – динамічне обчислення порогу для різних ділянок зображення.

У програмі реалізовано адаптивний поріг, оскільки фон номерного знака не завжди є рівномірно білим. Як базове порогове значення було обрано сірий колір, що дозволило ефективно відокремити символи (літери та цифри) від фону.

Процес сегментації:

- перетворення зображення в чорно-білий формат (градації сірого);
- застосування адаптивного порогу для виділення контурів символів;
- визначення контурів літер та цифр за допомогою функцій OpenCV;
- перетворення знайдених контурів у растрові зображення окремих символів;
- збереження кожного символу як окремого файлу у визначеній директорії для подальшого розпізнавання нейронною мережею.

Програмна реалізація процедури сегментації наведена у Додатку Б.

3.6 Архітектура нейронної мережі

Реалізований у роботі алгоритм розпізнавання символів ґрунтується на архітектурі нейронної мережі LeNet.

Схему мережі можна побачити на рисунку 3.3.

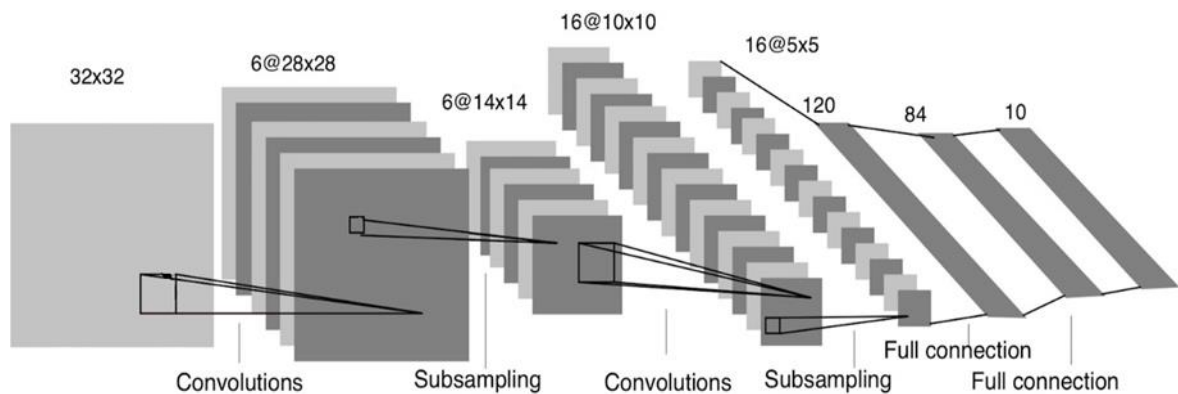


Рисунок 3.3 – Мережа LeNet

Для навчання моделі було використано генератори зображень з бібліотеки Keras, які забезпечили завантаження та попередню обробку даних із набору Train_Cells.

Аугментація даних. З метою підвищення узагальнюючої здатності моделі, окрім оригінальних зображень, було застосовано перетворені варіанти, що включали:

- зміну розміру зображень;
- випадкове масштабування окремих фрагментів;
- корекцію контрастності;
- регулювання яскравості.

Це дозволило розширити тренувальний набір та покращити стійкість мережі до варіацій у вхідних даних.

Архітектура згорткової нейронної мережі. Модель було побудовано з використанням Keras, відповідно до схеми (рисунок 3.3).

Основні характеристики архітектури:

а) Згорткові шари:

1) використано ініціалізацію ваг за нормальним розподілом.

б) повнозв'язні (Dense) шари:

1) перший та другий шари – функція активації ReLU (Rectified Linear Unit).

2) третій (вихідний) шар – функція Softmax (для багатокласової класифікації).

Оптимізація навчання

Для запобігання проблемам, пов'язаним із функціями активації («вмирання ReLU» та «зникаючий градієнт»), було обрано алгоритм оптимізації Adamax (розширена версія Adam).

Ключові особливості Adamax:

- використовує оцінку інерційного моменту (замість дисперсії, як у Adam);
- забезпечує адаптивне навчання з урахуванням градієнтів;
- реалізовано у Keras, що спрощує інтеграцію.

Процес навчання:

- тренування на оригінальних та аугментованих даних(100 епох);
- подавання сегментованих символів (після обробки знімка НЗ) на вхід мережі;
- класифікація кожного символу окремо.

Експериментально доведено, що підвищення контрастності зображень символів підвищує точність розпізнавання.

Програмна реалізація доступна у Додатку Б.

3.7 Приклад роботи алгоритму

Для тестування роботи алгоритму йому на вхід було подано знімок номерної пластини, представлений на рисунку 3.4.

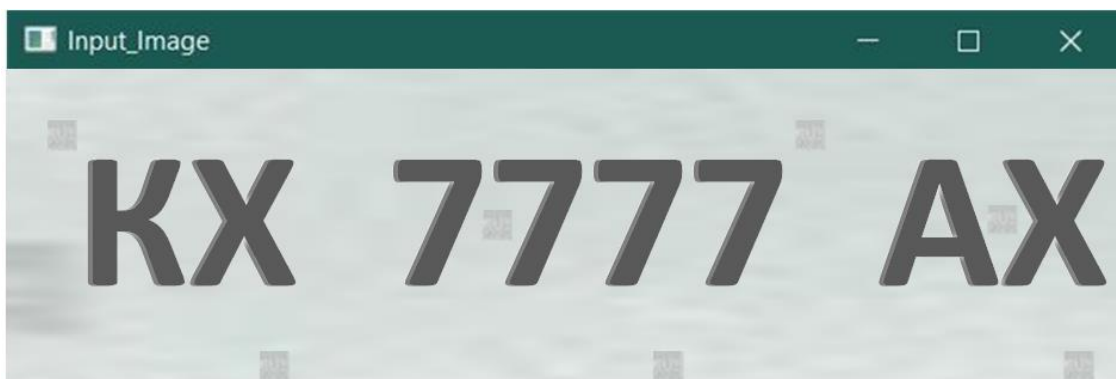


Рисунок 3.4 – Знімок номерного знака

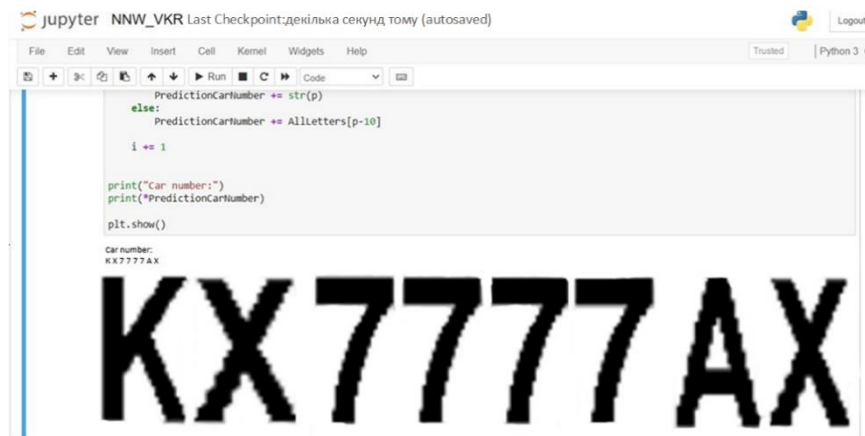
Алгоритм сегментації виділив такі символи на вхідному зображенні (рисунок 3.5).



Рисунок 3.5 – Контури символів

На рисунку 3.6 представлені зображення символів (з доданою

контрастністю та зміненням розміром) та результат роботи нейронної мережі – розпізнаний номер транспортного засобу.



```

PredictionCarNumber += str(p)
else:
    PredictionCarNumber += AllLetters[p-10]
    i += 1

print("Car number:")
print(PredictionCarNumber)
plt.show()

```

Car number:
KX7777 AX

Рисунок 3.6 – Результат роботи нейромережі

Як видно з рисунка 3.6, нейронна мережа успішно розпізнала всі символи на вхідному зображенні номерної пластини. Це свідчить про ефективність обраної архітектури згорткової мережі та якість попередньої обробки даних, включаючи адекватну сегментацію зображення та аугментацію тренувального набору. Використання оптимальних функцій активації (ReLU у прихованих шарах та Softmax на виході) разом із алгоритмом оптимізації Adamax забезпечило стабільне навчання моделі без типових проблем, пов'язаних із зникненням градієнтів.

Експериментальні результати підтверджують, що запропонований підхід до розпізнавання номерних знаків є практично придатним. Особливо позитивно на точність вплинуло підвищення контрастності зображень під час передобробки. Подальші дослідження могли б спрямуватись на покращення стійкості системи до складних умов, таких як низька якість знімків або нестандартні шрифти символів, що залишається актуальним напрямком для майбутніх робіт у цій галузі.

ВИСНОВКИ

У кваліфікаційній роботі було досліджено основні методи розпізнавання об'єктів на зображеннях, зокрема особливості проектування та навчання згорткових нейронних мереж (ЗНМ). Практична частина роботи включала реалізацію процедури сегментації зображень та створення спеціалізованої ЗНМ для розпізнавання символів на державних номерних знаках транспортних засобів. Для реалізації проекту було обрано мову програмування Python з використанням бібліотек Keras для роботи з нейронними мережами та OpenCV для обробки зображень.

В якості основного інструменту сегментації був застосований метод Thresholding, тоді як для розпізнавання символів використана оптимізована архітектура LeNet з алгоритмом оптимізації Adamax. Навчання мережі проводилось на вибірці з 20 000 зображень, що дозволило досягти високої точності розпізнавання. Отримані результати демонструють ефективність згорткових нейронних мереж для розв'язання вузькоспеціалізованих задач комп'ютерного зору, зокрема у сфері автоматичного розпізнавання номерних знаків. Висока продуктивність та адаптивність таких мереж відкривають широкі перспективи для їх застосування в різних галузях, де необхідна обробка та аналіз візуальних даних.

Використання згорткових нейронних мереж дозволяє будувати якісні алгоритми розпізнавання об'єктів на фотографіях та відеофайлах. Як правило, структуру ЗНМ підбирають відповідно до особливостей задачі та з вимогами до її вирішення. Орієнтація на вузькоспеціалізоване завдання дозволяє отримувати точні та коректні рішення. Також згорткові мережі відрізняються високою продуктивністю та ефективністю. Нині сфер їх застосування стає дедалі більше.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. За рік українці витратили на нові авто 125 млрд гривень [Електронний ресурс] // UkrAutoProm. – Режим доступу: <https://ukrautoprom.com.ua/za-rik-ukrayinczi-vytratyly-na-novi-avto-125-mlrd-gryven> (дата звернення: 17.05.2025).
2. У 2024 році автопарк України поповнили понад 2221 тис. вживаних легковиків [Електронний ресурс] // UkrAutoProm. – Режим доступу: <https://ukrautoprom.com.ua/u-2024-roczii-avtopark-ukrayiny-popovnyly-ponad-2221-tys-vzhyvanyh-legkovykyv> (дата звернення: 17.05.2025).
3. Боднар, А., Польшакова, О. Автоматизоване розпізнавання номеру транспортного засобу в режимі реального часу за допомогою технологій комп'ютерного зору / А. Боднар, О. Польшакова // Адаптивні системи автоматичного управління. – 2025. – № 46. – С. 17–24. – DOI: 10.20535/1560-8956.46.2025.323693.
4. Вовк, С. М., Гнатушенко, В. В., Бондаренко, М. В. Методи обробки зображень та комп'ютерний зір : навч. посіб. / С. М. Вовк, В. В. Гнатушенко, М. В. Бондаренко. – Д. : ЛІРА, 2016. – 148 с. – ISBN 978-966-383-699-7.
5. Бондаренко, М. В. Теорія графів і мереж : навч. посіб. / М. В. Бондаренко. – К. : . 2004. – 480 с.
6. Philips Precise Image: AI for significantly lower dose and improved image quality [Електронний ресурс] // Philips. – Режим доступу: <https://www.documents.philips.com/assets/20240223/4f2cbfa4d0df428f94bfb11f01751a87.pdf> (дата звернення: 19.05.2025).
7. All the new generative AI camera features in Google Pixel 9 phones [Електронний ресурс] // Wired. – Режим доступу: <https://www.wired.com/story/all-the-new-generative-ai-camera-features-in-google-pixel-9-phones/> (дата звернення: 19.05.2025).
8. Розпізнавання номерів з камер відеоспостереження [Електронний

ресурс]. – Режим доступу: <https://kobi.ua/solutions/rozpiznavannya-nomeriv-cameras/> (дата звернення: 19.05.2025).

9. Модулі аналітики – INNI [Електронний ресурс]. – Режим доступу: <https://www.inni.tech/moduli-analitiki> (дата звернення: 19.05.2025).

10. Vehicle License Plate Recognition – VCloud.AI [Електронний ресурс]. – Режим доступу: <https://ua.vcloud.ai/vehicle-license-plate-recognition/> (дата звернення: 19.05.2025).

11. Розпізнавання номерів камерою [Електронний ресурс] // Kobi.ua. – Режим доступу: <https://kobi.ua/solutions/rozpiznavannya-nomeriv-cameras/> (дата звернення: 21.05.2025).

12. Модулі аналітики [Електронний ресурс] // Inni.tech. – Режим доступу: <https://www.inni.tech/moduli-analitiki> (дата звернення: 19.05.2025).

13. Vehicle license plate recognition [Електронний ресурс] // VCloud.ai. – Режим доступу: <https://ua.vcloud.ai/vehicle-license-plate-recognition/> (дата звернення: 19.05.2025).

14. Про затвердження зразків, технічного опису та Порядку виготовлення, зберігання і видачі номерних знаків : наказ Міністерства внутрішніх справ України № 166 від 26 лютого 2009 р. [Електронний ресурс]. – Режим доступу: <https://zakon.rada.gov.ua/laws/show/z0276-09#Text> (дата звернення: 05.06.2025).