МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ЗАПОРІЗЬКА ПОЛІТЕХНІКА»

Факультет комп'ютерних наук та технологій
Кафедра комп'ютерних систем та мереж

Пояснювальна записка
до дипломного проєкту (роботи)
бакалавра
(ступінь вищої освіти (освітній ступінь))

на тему «РОЗРОБКА КОМП'ЮТЕРНОЇ СИСТЕМИ РОЗПІЗНАВАННЯ ГРАФІЧНИХ ЗОБРАЖЕНЬ НА БАЗІ НЕЙРОННИХ МЕРЕЖ»

Виконав: студент 4 курсу, групи КНТ-512сп
спеціальності 123 Комп'ютерна інженерія

Освітня програма (спеціалізація)
Комп'ютерна інженерія
(код і назва спеціальності)

НІКІТЮК А.І.
(прізвище та ініціали)
Керівник ІЛЬЯШЕНКО М.Б.
(прізвище та ініціали)
Рецензент КОЗІНА Г.Л.
(прізвище та ініціали)

2025

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Національний університет «Запорізька політехніка»

(повне найменування вищого навчального закладу)

Факультет Комп'ютерних наук і технологій

Кафедра «Комп'ютерні системи та мережі»

Ступінь вищої освіти (освітній ступінь) бакалаврський

Спеціальність 123 Комп'ютерна інженерія

(код і назва)

Освітня програма (спеціалізація) Комп'ютерна інженерія

(назва)

ЗАТВЕРДЖУЮ
Зав. кафедри
_____ Кудерметов Р.К.
"____" _____ 2025 року

З А В Д А Н Н Я
НА ДИПЛОМНИЙ ПРОЄКТ СТУДЕНТУ

НІКІТЮК Антон Ігоревич

(прізвище, ім'я, по батькові)

1. Тема проекту (роботи) «Розробка комп'ютерної системи розпізнавання графічних зображень на базі нейронних мереж»

керівник проекту (роботи) ІЛЬЯШЕНКО М.Б. к. т. н., доцент

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом вищого навчального закладу від "08" квітня 2025 року № 151

2. Строк подання студентом проекту (роботи) 01.06. 2025року

3. Вихідні дані до проекту (роботи) нейромережі, мова програмування, база даних

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити)

1) Опис предметної області

2) Реалізація класифікації об'єктів при розпізнаванні

3) Проєктування системи розпізнавання

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень)

Слайди презентації

6. Консультанти розділів проекту (роботи)

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	прийняв виконане завдання
1-3	ІЛЬЯШЕНКО М.Б. к.т.н.. доцент		
Нормоконтроль	ЩЕРБАК Н.В, ст. викладач		

7. Дата видачі завдання 01.04. 2025 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломного проекту (роботи)	Строк виконання етапів проекту (роботи)	Примітка
1	Опис предметної області	05.04.2025	
2	Принцип роботи нейронних мереж	10.05.2025	
3	Вибір типа мережі	17.04.2025	
4	Вимоги до проєкту	23.04.2025	
5	Розробка системи	18.05.2025	
6	Тестування системи	25.05.2025	
7	Оформлення отриманих результатів у ПЗ	26.05.2025	
8	Оформлення графічного матеріалу	29.05.2025	
9	Захист роботи		

Студент _____ **Антон НІКІТЮК**
(підпис) (ініціали та прізвище)

Керівник проекту (роботи) _____ **Матвій ІЛЬЯШЕНКО**
(підпис) (ініціали та прізвище)

РЕФЕРАТ

Пояснювальна записка до дипломної кваліфікаційної роботи бакалавра: 65 с., 26 рис., 1 табл., 15 джерел.

АЛГОРИТМІЧНІ ПОБУДОВИ, ГРАФІЧНИ ОБРАЗИ, НЕЙРОННІ МЕРЕЖІ, РЕЦЕПТОРНА СТРУКТУРА

Мета роботи - методи розпізнавання графічних образів.

Методи дослідження - розробка, проектування і тестування нейронної мережі для полегшення процедури розпізнавання графічних образів.

Результати роботи - полягають у експериментальних дослідженнях методів розпізнавання графічних образів з використанням нейронних мереж. Розробка комп'ютерної системи розпізнавання графічних образів.

Рекомендації по використанню результатів роботи - отримані результати досліджень можна використати для розпізнавання різних типів графічних образів.

Можливі напрямки розвитку - створення програмних бібліотек для різних методів розпізнавання графічних образів використовуючи нейронні мережі.

ABSTRACT

Explanatory note to the master’s work: 65 p., 26 figures, 1 tables, 15 sources.

ALGORITHMIC CONSTRUCTIONS, GRAPHICAL IMAGES, NEURAL NETWORKS, RECEPTOR STRUCTURE

Purpose of the work - ethods of graphic image recognition.

Research methods - development, design and testing of a neural network to facilitate the procedure of graphic image recognition.

Results of the work - experimental studies of graphic image recognition methods using neural networks. Development of a computer system for recognizing graphic images.

Recommendations for the use of the results of the work - the obtained research results can be used to recognize different types of graphic images.

Possible directions of development - creation of software libraries for various methods of graphic image recognition using neural networks.

ЗМІСТ

Вступ 7
1 Опис предметної області 9
1.1 Опис існуючих реалізацій 9
1.2 Огляд існуючих реалізацій 12
1.3 Нейронні мережі для вирішення задачі 20
1.3.1 Застосування нейронних мереж 23
1.3.2 Переваги нейронних мереж 23
2 Реалізація класифікаціяї об’єктів при розпізнаванні 24
2.1 Структура сприйняття інформації при розпізнаванні 24
2.2 Механізм побудови нейромережі 35
2.2.1 Методи навчання нейромереж 36
2.3 Засоби розпізнавання 44
3 Проєктування системи розпізнавання 48
3.1 Порядок дій при реалізації системи 48
3.1.1 Алгоритм нейромережевої організації матриць для розпізнавання образів 50
3.2 Реалізація системи розпізнавання 54
3.3 Аналіз результатів 59
Висновки 62
Перелік джерел посилання 64

ВСТУП

Актуальність теми. Зростаючий інтерес до задач розпізнавання образів пояснюється необхідністю автоматизації як функцій контролю та управління складними динамічними об'єктами в реальному часі, так і процесів комунікації в інтелектуальних системах. Тому досі триває пошук і впровадження ефективних методів передачі функцій розпізнавання людини комп'ютеризованим системам. Одним із перспективних підходів до вирішення цієї проблеми є використання штучних нейронних мереж і нейрокомп’ютерів, які найбільш придатні для розв’язання задач розпізнавання образів.

Сьогодні існує велика кількість нейромережевих підходів до розпізнавання образів. Однак складності виникають у випадках, коли зображення піддаються спотворенням, таким як зашумлення, зміщення, поворот або зміна масштабу. Цю проблему можна вирішити шляхом вибору відповідної архітектури нейромережі та методу навчання. Проте аналіз наукових робіт показує, що наразі не існує універсальної моделі, нечутливої до всіх типів спотворень.

Задачу розпізнавання зображень, що зазнали зміщення або зашумлення, ефективно вирішують нейронні мережі зі зворотним розповсюдженням помилки. Однак досі залишаються труднощі у випадках зміни масштабу чи повороту об'єкта. Перспективним напрямом у подоланні цих проблем є використання нової нейромережевої парадигми – моделі персептрона, яка має принципово іншу архітектуру та використовує неконтрольоване навчання. Архітектура персептрона базується на організації зорової системи людини.

Мета роботи – розробка, проєктування та тестування вдосконаленої нейронної мережі для спрощення процесу розпізнавання образів.

Методи дослідження**.** У роботі розглянуто різні підходи до вирішення задачі розпізнавання образів, зосереджено увагу на застосуванні штучних нейронних мереж та методах їх навчання. Особливу увагу приділено аналізу персептрона – його переваг та недоліків. У завершальній частині роботи

розроблено та запропоновано вдосконалену модель персептрона, наведено результати її реалізації та порівняно з прототипом, що дозволило оцінити переваги та недоліки запропонованого підходу.

Наукова новизна. Вдосконалено метод розпізнавання графічних образів за допомогою нейронної мережі, що дозволило підвищити швидкість обробки, зменшити апаратні витрати та ефективно розпізнавати специфічні позначки на зображеннях.

Практична цінність. Результати роботи можуть бути використані для ефективного застосування нейронних мереж у задачах розпізнавання образів. Крім того, запропоновану модель можна адаптувати та вдосконалювати для досягнення ще кращих результатів.

1 ОПИС ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Опис існуючих реалізацій

Довгий час розпізнавання графічних образів розглядалося виключно з біологічної та психологічної точки зору. В основному вивчалися якісні характеристики процесу, що не дозволяло чітко описати механізми його функціонування. Дослідження здебільшого зосереджувалися на рецепторах органів чуття (зору, слуху, дотику), проте принципи прийняття рішень залишалися невідомими. На початкових етапах помилково вважалося, що мозок працює за жорстко визначеними алгоритмами, які можна відтворити технічними засобами.

На початку XX століття Норберт Вінер започаткував нову науку – кібернетику, яка вивчає загальні закономірності управління та передачі інформації в живих організмах, технічних пристроях і суспільстві. Це дозволило перейти від якісного аналізу розпізнавання образів до кількісних методів, тобто застосувати математичний апарат для опису природних процесів розпізнавання.

Розробка пристроїв, що виконують функції розпізнавання, дає змогу замінити людину автоматизованими системами, розширюючи можливості складних інформаційних, логічних та аналітичних процесів. Людська робота залежить від багатьох чинників (досвіду, кваліфікації, уважності), тоді як справний автомат працює стабільно та забезпечує незмінну якість. Автоматизований контроль дозволяє здійснювати моніторинг складних систем, забезпечувати своєчасне обслуговування, ідентифікувати перешкоди та використовувати методи шумозаглушення, що покращує передачу інформації. Крім того, автоматизовані системи здатні виконувати завдання з високою швидкістю, недосяжною для людини [1].

Проблема розпізнавання образів давно привертає увагу фахівців у галузі прикладної математики та інформатики. Ще у 1920-х роках Р. Фішер розробив методи, що заклали основу дискримінантного аналізу – важливого розділу

розпізнавання. У 1940-х роках А. М. Колмогоров та А. Я. Хінчин вивчали розподіл сумішей випадкових величин.

У 1950-1960-х роках з'явилася теорія статистичних рішень, що дала змогу створити алгоритми класифікації об'єктів за наперед визначеними класами. Це стало початком цілеспрямованих наукових досліджень та практичних розробок. В межах кібернетики сформувався новий науковий напрям, присвячений теоретичним основам та практичному застосуванню систем розпізнавання об'єктів, процесів і явищ, який отримав назву "Розпізнавання образів".

Таким чином, основою для класифікації об'єктів стали результати класичної теорії статистичних рішень. У її межах розроблялися алгоритми, що дозволяли визначати клас об'єкта на основі вимірюваних параметрів (ознак) та деяких апріорних знань про класи.

Розпізнавання образів – це процес ідентифікації об’єкта або визначення його характеристик за допомогою оптичного (зображення), акустичного (звуки) чи інших методів (рис. 1.1).

Образ – це класифікаційна категорія, що об'єднує певну групу об'єктів за спільною ознакою. Характерна властивість образів полягає в тому, що знайомство з обмеженою кількістю їхніх представників дає змогу розпізнавати необмежену кількість нових об'єктів тієї ж групи.

Методика визначення відповідності об'єкта певному класу називається вирішальним правилом. Важливим поняттям у розпізнаванні є метрика – спосіб визначення відстані між об'єктами у множині. Чим менша ця відстань, тим подібнішими є об'єкти (наприклад, символи, звуки тощо).

Зазвичай об'єкти подаються у вигляді набору чисел, а метрика – у вигляді функції. Вибір способу представлення образів і метрики впливає на ефективність розпізнавання: один і той самий алгоритм може давати різні результати залежно від використаної метрики.

Адаптація - це процес коригування параметрів, структури системи та, за необхідності, керуючих впливів на основі актуальної інформації. Основна мета адаптації - досягнення бажаного стану системи в умовах початкової

невизначеності та змінного середовища.

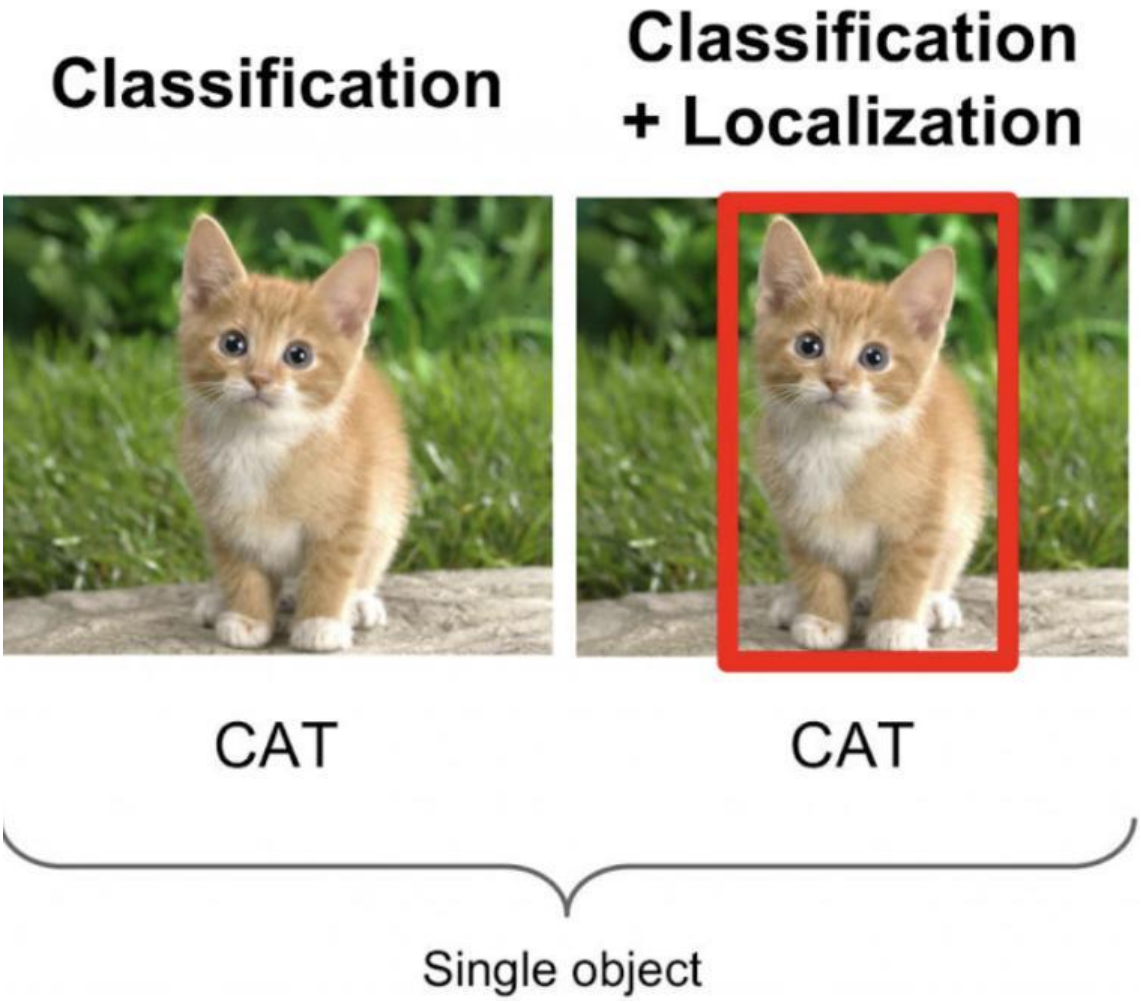


Рисунок 1.1 – Приклад розпізнавання даних на фото

Навчання - це процес, у ході якого система поступово набуває здатності правильно реагувати на певні комбінації зовнішніх впливів. Водночас адаптація передбачає налаштування параметрів і структури системи для забезпечення необхідної якості управління в умовах постійної зміни зовнішніх факторів [2].

Навчанням зазвичай називають процес формування у системи певної реакції на повторювані зовнішні сигнали через багаторазову корекцію з боку зовнішнього впливу. Таку корекцію прийнято позначати як «заохочення» та «покарання». Алгоритм навчання значною мірою визначається механізмом формування цих коригувальних сигналів.

Самонавчання відрізняється тим, що система не отримує додаткової

інформації про правильність власних реакцій, а формує її самостійно, спираючись на внутрішні механізми аналізу та оцінки своїх рішень.

1.2 Огляд існуючих реалізацій

Методи розпізнавання графічних образів поділяються на три основні категорії: попередня фільтрація та підготовка зображення, логічна обробка отриманих даних і алгоритми ухвалення рішень на основі цієї обробки. Важливо зазначити, що межі між цими категоріями умовні, і для розв’язання конкретного завдання не завжди необхідно використовувати всі три методи — іноді достатньо лише одного або двох.

Ця категорія включає методи, які допомагають виділити певні області на зображенні без їхнього безпосереднього аналізу. Більшість таких методів застосовують однакове перетворення для всіх точок зображення. На цьому етапі зображення ще не аналізується, проте точки, що проходять фільтрацію, можуть розглядатися як області з особливими характеристиками.

Одним із найпростіших методів є бінаризація зображення, що передбачає перетворення кольорового або відтінкового зображення в двійковий формат на основі певного порогового значення. У деяких ідеальних випадках цього перетворення достатньо для вирішення задачі. Наприклад, якщо потрібно автоматично виокремити об’єкти на білому фоні, вибір відповідного порогу визначатиме якість бінаризації. Зазвичай використовується адаптивний алгоритм, який автоматично встановлює оптимальний поріг [3].

Бінаризація також може бути ефективною при роботі з гістограмами або під час аналізу зображень у просторі HSV, а не RGB. Це дозволяє здійснювати, наприклад, сегментацію кольорів, що є основою для створення детекторів міток або розпізнавання кольору людської шкіри.

До традиційних методів фільтрації належать перетворення Фур’є,

низькочастотні фільтри (ФНЧ) та високочастотні фільтри (ФВЧ). Ці методи, запозичені з радіолокації та обробки сигналів, широко застосовуються у сфері розпізнавання образів.

Перетворення Фур'є, зокрема його швидка версія (БПФ), рідко використовується в аналізі зображень у чистому вигляді. Проте воно знаходить застосування у стисненні зображень [4]. Для більш глибокого аналізу зазвичай необхідне двовимірне перетворення Фур'є (1.1), хоча воно є значно більш ресурсомістким, ніж його одновимірний аналог:

$$G_{UM} = \frac{1}{NM} \sum_{n=1}^{N-1} \sum_{m=1}^{M-1} x_{mn} e^{-2\pi ij} \tag{1.1}$$

Лише одиниці займаються безпосереднім розрахунком перетворення Фур'є, оскільки значно швидше та зручніше застосовувати згортку області з готовим фільтром, налаштованим на високі (ФВЧ) або низькі (ФНЧ) частоти. Хоча цей метод не дозволяє проводити спектральний аналіз, у задачах відеоспостереження зазвичай важливий не сам аналіз, а отримання конкретного результату.

До найпростіших прикладів фільтрів, що підкреслюють певні частоти, належать фільтр Калмана (для низьких частот) та фільтр Габора (для високих частот). Їхня реалізація передбачає вибір вікна для кожної точки зображення та його згортку з фільтром аналогічного розміру. У результаті цього процесу кожна точка зображення отримує нове значення. При застосуванні ФНЧ та ФВЧ виходять зображення відповідного типу.

Але що станеться, якщо замість стандартного фільтра використати для згортки довільну характеристичну функцію? У такому випадку цей процес називається вейвлет-перетворенням. Хоча це визначення не є строго математично коректним, у багатьох дослідницьких командах під вейвлет-аналізом розуміють пошук заданого патерну на зображенні шляхом згортки з його моделлю.

У класичному вейвлет-аналізі використовуються такі відомі функції, як

вейвлет Хаара, вейвлет Морле та "мексиканський капелюх". Вони є основою багатьох методів обробки сигналів та зображень.

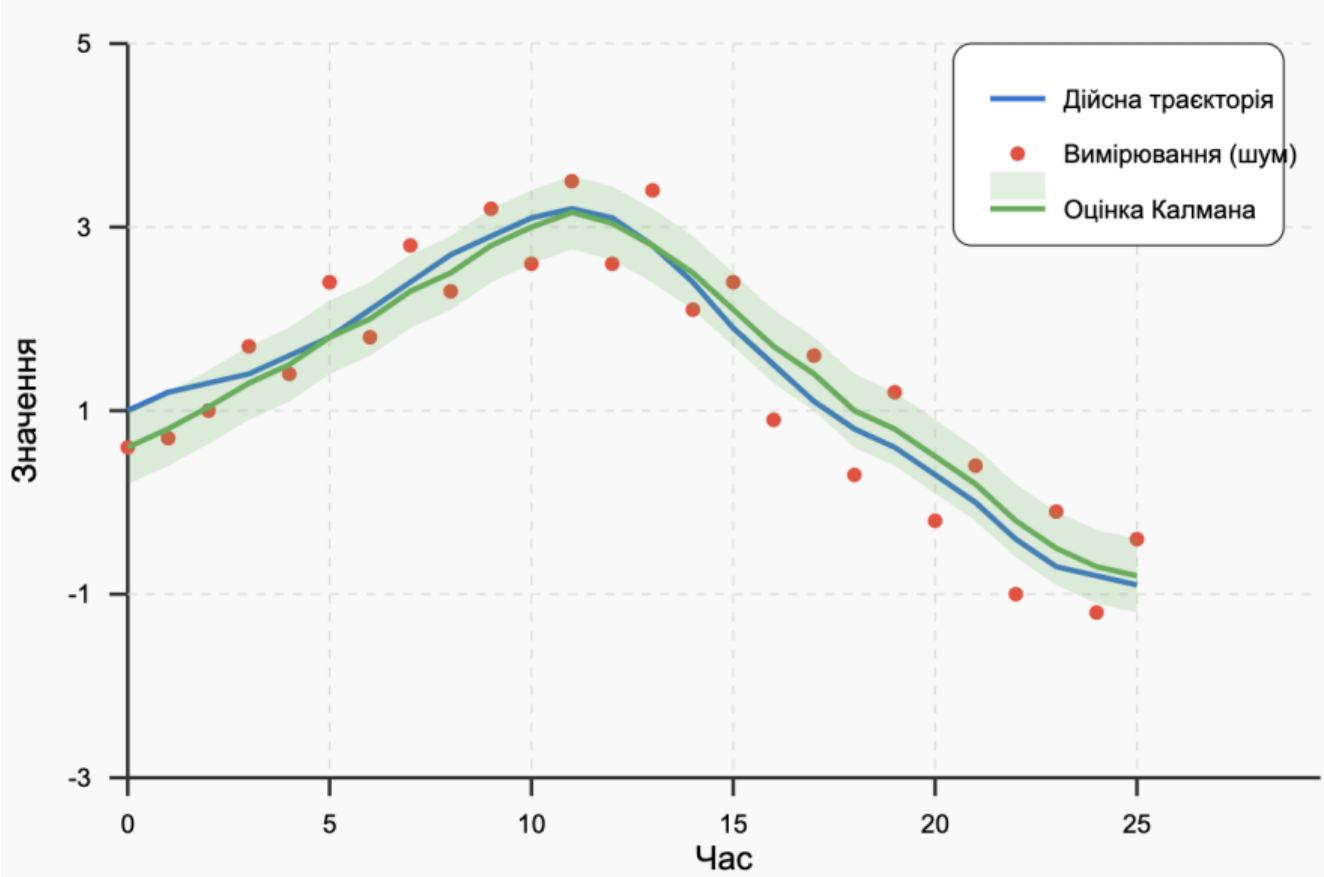


Рисунок 1.2 – Фільтр Калмана

До класичних прикладів вейвлетів належать: тривимірний вейвлет Хаара, двовимірні вейвлети Мейєра, Мексиканська Капелюх і Добеші.

Розширене трактування вейвлет-аналізу знаходить практичне застосування в різних задачах, наприклад, у виявленні відблисків в оці, де сам відблиск виступає в ролі вейвлету.

Класичні вейвлети найчастіше використовуються у стисненні зображень або для їх класифікації, що буде розглянуто далі (1.2, рис.1.3):

$$f(x) = \exp\left(-\frac{x^2}{2\sigma^2}\right)\cos(2\pi\theta x), \qquad (1.2)$$

де $\theta = 1$;

$\sigma=3,01$.

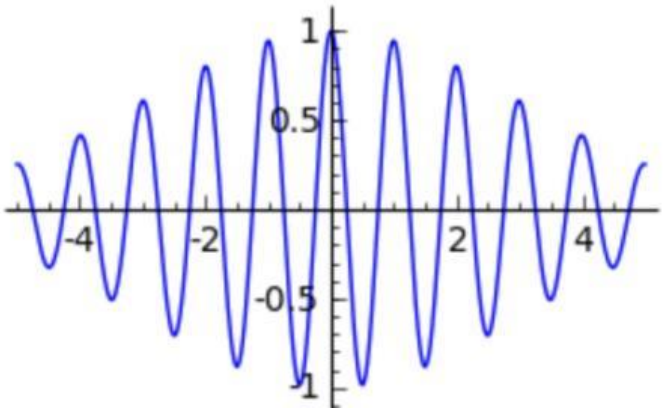


Рисунок 1.3 – Фильтр Гарбора

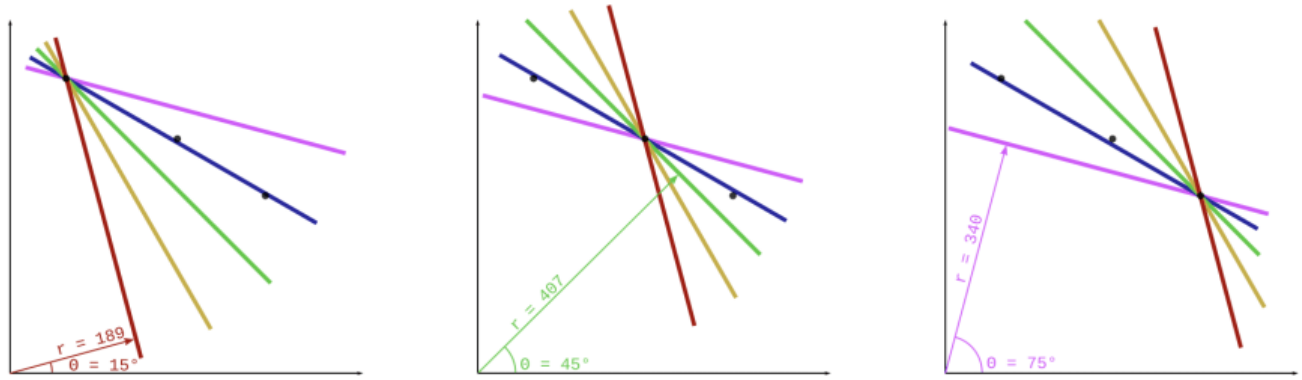
Після такої вільної інтерпретації вейвлетів варто згадати кореляцію, яка лежить в їх основі. У фільтрації зображень цей метод є незамінним інструментом.

Класичне застосування кореляції - аналіз відеопотоку для виявлення зміщень або оптичних потоків. Найпростіший детектор зміщення, по суті, є різницевим корелятором: у тих зонах, де зображення не корелюють, відбувався рух.

Цікавим типом фільтрації є фільтрація функцій. Це математичні фільтри, які дозволяють знаходити просту геометричну фігуру на зображенні (пряму, параболу, коло тощо). Методика передбачає побудову нового зображення, де для кожної точки відзначається множина функцій, які могли її породити.

Класичним прикладом цього підходу є перетворення Хафа для визначення прямих (рис. 1.4). У ньому кожна точка (x, y) відображається у простір параметрів (a, b) відповідної прямої $y = ax + b$, яка проходить через цю точку.

Перетворення Хафа дозволяє знаходити будь-які параметризовані функції, включаючи окружності. Існує його модифікація, яка дає змогу виявляти довільні фігури, що дуже цінують математики. Однак, при обробці зображень цей метод має значні обмеження: він працює повільно та чутливий до якості бінаризації. Навіть у найкращих умовах я зазвичай віддаю перевагу іншим методам.



θ	r
15	189.0
30	282.0
45	355.7
60	407.3
75	429.4

θ	r
15	318.5
30	376.8
45	407.3
60	409.8
75	385.3

θ	r
15	419.0
30	443.6
45	438.4
60	402.9
75	340.1

Рисунок 1.4 – Перетворення Хафа для прямих

Аналогом перетворення Хафа для прямих є перетворення Радона. Його обчислення базується на швидкому перетворенні Фур'є (БПФ), що забезпечує виграш у продуктивності, особливо якщо точок дуже багато. До того ж, цей метод можна застосовувати до небінаризованих зображень.

Фільтрація контурів та кордонів. Контури є ключовими для переходу від обробки зображень до аналізу об'єктів. Коли об’єкт складний, але добре виділений, часто єдиним способом його обробки стає виділення контурів.

Серед алгоритмів фільтрації контурів найпопулярнішим є детектор Кенні, який ефективний і має реалізацію в OpenCV. Також існує оператор Собеля, але він гірше справляється із завданням пошуку контурів.

Логічна обробка результатів фільтрації. Фільтрація зображень дає набір даних, придатних для подальшої обробки. Проте часто отримані результати потребують додаткової математичної інтерпретації.

Проміжним кроком між фільтрацією та логічною обробкою є методи математичної морфології. Це базові операції ерозії та дилатації бінарних зображень, які допомагають прибрати шуми, змінюючи розмір наявних елементів. На основі цих методів будуються алгоритми виділення контурів, хоча зазвичай використовуються гібридні підходи.

Раніше згадувалися алгоритми отримання кордонів. Отримані кордони легко перетворюються на контури: для детектора Кенні це відбувається автоматично, а для інших методів потрібна додаткова бінаризація [5].

Контур є унікальною характеристикою об'єкта, що дозволяє його ідентифікувати. Для цього існує контурний аналіз—потужний математичний апарат, який дозволяє розпізнавати об'єкти за їхніми контурами.

Особливі точки об'єкта - мце унікальні характеристики, які дозволяють зіставляти об'єкти між собою або з іншими класами об'єктів. Існують десятки методів для їхнього визначення.

Деякі алгоритми знаходять особливі точки у сусідніх кадрах відео, інші- при значних змінах освітлення або поворотах об'єкта. Спершу розглянемо швидкі методи, що знаходять менш стабільні особливі точки, а потім перейдемо до більш складних, але надійних.

Перший клас - це нестабільні особливі точки, які зберігаються протягом секунд. Вони використовуються для відстеження об'єкта у сусідніх кадрах або для зіставлення зображень із різних камер.

До таких точок належать:

- локальні максимуми яскравості;
- кути на зображенні (рис. 1.5);
- точки з максимальною дисперсією;
- градієнти та їхні екстремуми.

Ці методи забезпечують швидкий розрахунок, але не завжди достатньо надійні для складних задач.

Другий клас - стабільні особливі точки.

Ці точки залишаються незмінними при зміні освітлення і невеликих рухах об'єкта. Вони широко застосовуються у навчанні моделей та класифікації об'єктів.

Прикладом використання таких точок є:

- класифікатор пішоходів;
- системи розпізнавання облич.

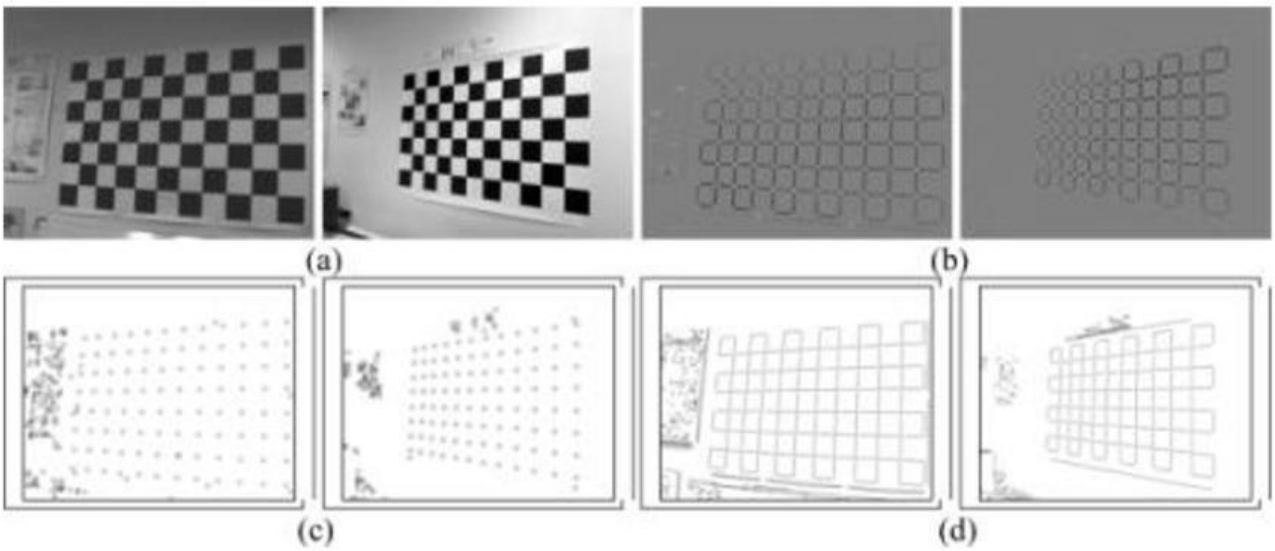


Рисунок 1.5 – Розпізнавання кутів на зображеннях

Частина вейвлет-перетворень, розглянутих раніше, може бути основою для визначення подібних точок. Наприклад:

- примітиви Хаара;
- пошук відблисків;
- аналіз специфічних функцій.

Одним із найефективніших методів знаходження стабільних особливих точок є гістограми спрямованих градієнтів (HOG) (рис. 1.6).

Третій клас. Повністю стабільні особливі точки. Є лише два методи, що забезпечують повну стабільність особливих точок, разом із їхніми модифікаціями:

- SIFT (Scale-Invariant Feature Transform);
- SURF (Speeded-Up Robust Features).

Ці методи дозволяють виявляти особливі точки навіть при зміні масштабу та поворотах зображення. Вони обчислюються довше у порівнянні з іншими підходами, але загалом їхній розрахунок займає досить обмежений час. Недолік: обидва методи запатентовані, що обмежує їхнє використання в комерційних проектах.

Навчання та прийняття рішень. Ця частина присвячена методам, які не працюють безпосередньо з зображенням, але використовуються для аналізу та класифікації об'єктів. Зокрема, це методи машинного навчання.

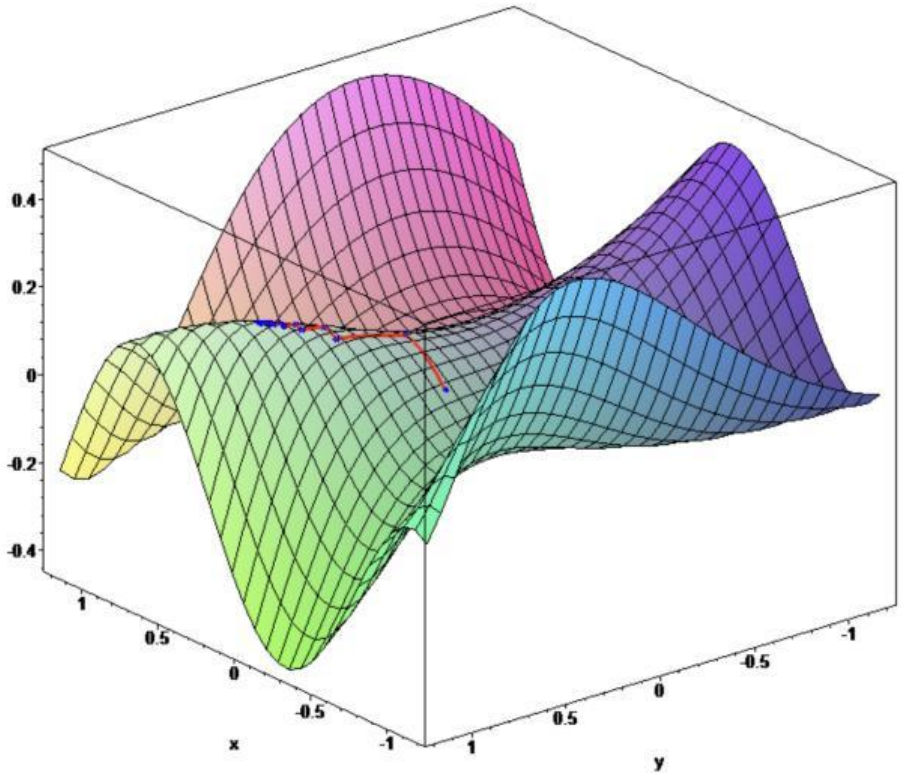


Рисунок 1.6 – Гістограм спрямованих градієнтів

Приклад навчання. Припустимо, ми створюємо класифікатор, що визначає наявність людини на фото. Кожне зображення представлено у вигляді набору ознак (наприклад, HOG, Haar, SURF, вейвлети).

Ознаковий простір для класифікації може містити:

- "Наявність очей";
- "Наявність носа";
- "Наявність двох рук";
- "Наявність вух".

Якщо всі ці ознаки виявлені, то отримуємо точку [1;1;1;1] – людина. Для мавпи – [1;0;1;0], для коня – [1;0;0;0].

Проте, через похибки детекторів або особливості фото деякі ознаки можуть бути відсутніми чи спотвореними. Класифікатор повинен автоматично розділити простір ознак, щоб чітко визначати об’єкти.

Як працює класифікація? Кожне зображення – це точка в просторі ознак, а класифікатор повинен знайти кордони між класами об'єктів.

Наприклад, один із популярних алгоритмів – AdaBoost – поступово уточнює розподіл меж між класами для більш точної ідентифікації.

Вибір класифікатора. Класифікаторів існує багато, і кожен має свою оптимальну сферу застосування. Підбір класифікатора для конкретного завдання – це окрема наука і певною мірою мистецтво.

1.3 Нейронні мережі для вирішення задачі

Дослідження штучних нейронних мереж пов’язане з їхньою здатністю наближатися до принципів обробки інформації, характерних для людського мозку. Мозок – це складна, нелінійна, паралельна система, яка обробляє інформацію з високою ефективністю. Його унікальна особливість полягає у здатності самоорганізовувати нейрони таким чином, щоб виконувати конкретні завдання, як-от розпізнавання образів, обробка сенсорних сигналів чи управління рухами, значно швидше за найпотужніші сучасні комп’ютери (рис.1.7).

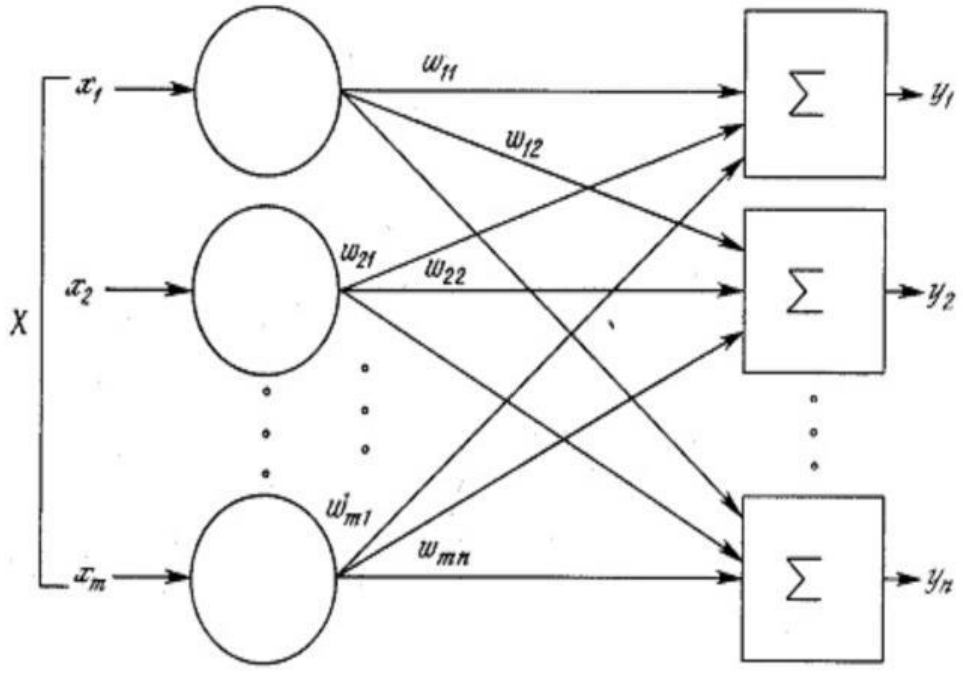


Рисунок 1.7 – Найпростіша одношарова нейронна мережа

Сьогодні штучні нейронні мережі широко застосовуються в різних сферах, зокрема у прогнозуванні, класифікації, оптимізації, розпізнаванні образів тощо.

Нейронні мережі – це обчислювальні моделі, що імітують певні біологічні процеси, характерні для людського мозку. Вони здатні до навчання на основі аналізу позитивних і негативних впливів. Основним елементом таких мереж є штучний нейрон, який названий так за аналогією з його біологічним прототипом.

Прототипом штучного нейрона став біологічний нейрон головного мозку. Він складається з тіла, дендритів – відростків, що приймають вхідні сигнали, і аксона, який передає вихідні сигнали іншим клітинам. Місце з'єднання аксона з дендритом називається синапсом [6].

Функціонування біологічного нейрона можна описати так (рис.1.8):

- отримання сигналів – нейрон приймає набір вхідних сигналів через дендрити;

- обробка сигналів – у тілі нейрона оцінюється сумарна величина отриманих сигналів. Проте різні входи мають різну значущість, яка визначається ваговими коефіцієнтами. Тому нейрон не просто підсумовує сигнали, а обчислює скалярний добуток вектора вхідних сигналів і відповідного вектора ваг;

- формування вихідного сигналу – якщо отримане значення перевищує певний поріг, нейрон «спрацьовує» і генерує вихідний сигнал. Якщо ж ні, сигнал не формується;

- передача сигналу – вихідний сигнал поширюється через аксон і передається іншим нейронам.

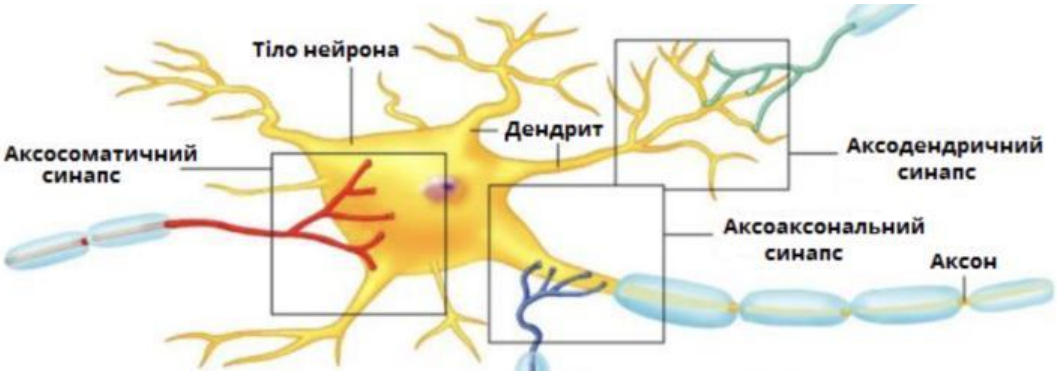


Рисунок 1.8 – Біологічний нейрон

Однією з ключових особливостей нейронних мереж є їх здатність навчатися на основі отриманих даних і з часом покращувати продуктивність. Це відбувається за рахунок коригування синаптичних ваг і порогів згідно з певними правилами навчання.

Навчання – складний процес, який можна визначити по-різному залежно від контексту. У випадку нейронних мереж можна розглядати його як налаштування параметрів системи на основі моделювання зовнішнього середовища.

Процес навчання передбачає такі етапи:

- надходження зовнішніх стимулів – нейронна мережа отримує вхідні дані;

- зміна параметрів – під впливом отриманих сигналів коригуються вагові коефіцієнти та інші внутрішні параметри мережі;

- адаптація поведінки – після оновлення структури мережа починає по-іншому реагувати на нові вхідні сигнали.

Окреслений набір чітких правил, за якими відбувається навчання нейронної мережі, називають алгоритмом навчання. Універсального алгоритму, що підходив би для всіх архітектур нейронних мереж, не існує. Натомість є безліч алгоритмів, кожен із яких має свої переваги. Вони відрізняються способом налаштування синаптичних ваг нейронів, а також типом взаємодії мережі із зовнішнім світом. У цьому контексті виділяють три основні парадигми навчання: з учителем, без учителя та змішане.

Навчання з учителем передбачає наявність правильних відповідей (цільових вихідних значень) для кожного вхідного вектора. Під час навчання ваги нейронів коригуються, поки похибка між отриманими та цільовими виходами не досягне прийнятного рівня. У посиленому варіанті навчання відома лише загальна оцінка правильності відповіді, а не саме значення виходу.

Навчання без учителя є більш наближеною до біологічної моделі. У цьому випадку навчальна вибірка містить лише вхідні вектори, а алгоритм навчання коригує ваги мережі так, щоб подібні входи давали подібні виходи. Це дозволяє виявляти приховані закономірності в даних та групувати зразки за спільними ознаками.

Змішане навчання поєднує обидва підходи: частина ваг коригується під керівництвом учителя, а частина – за допомогою самонавчання.

1.3.1 Застосування нейронних мереж

Класифікація образів – визначення, до якого класу належить вхідний сигнал (наприклад, рукописний символ або мовний сигнал). Використовується у розпізнаванні тексту, мови, медичних даних (наприклад, електрокардіограм), аналізі зображень [7].

Кластеризація (категоризація) – розподіл об’єктів на групи без попереднього визначення їхніх класів. Такий метод корисний у дослідженні властивостей даних, стисненні інформації та пошуку подібних об’єктів.

Апроксимація функцій – знаходження оцінки невідомої функції за обмеженою вибіркою вхідних і вихідних даних. Ця задача важлива в наукових дослідженнях та інженерних розрахунках.

Прогнозування – передбачення майбутніх значень на основі попередніх даних. Використовується у фінансовій сфері (прогнозування цін на біржі), метеорології та інших галузях.

Оптимізація – знаходження найкращого рішення у межах заданих обмежень. Прикладом є задача комівояжера, а також численні проблеми у математиці, техніці, економіці та медицині.

Асоціативна пам’ять – запам’ятовування та відтворення інформації на основі її змісту. На відміну від класичних обчислювальних систем, де доступ до пам’яті здійснюється за адресою, асоціативна пам’ять дозволяє знаходити потрібні дані навіть за частковим запитом. Це ефективний підхід для мультимедійних баз даних.

1.3.2 Переваги нейронних мереж

Адаптивність – нейронні мережі можуть змінювати свої параметри у відповідь на зміну умов навколишнього середовища. Вони здатні швидко перенавчатися, що корисно, наприклад, у фінансових системах, де дані змінюються в реальному часі. Однак надмірна адаптивність може призводити до втрати стійкості, якщо система реагує на несуттєві зміни [8].

Висока швидкодія – завдяки паралельній обробці інформації нейронні мережі можуть виконувати обчислення значно швидше, ніж традиційні алгоритми.

Відмовостійкість – при пошкодженні окремих нейронів або зв’язків продуктивність мережі знижується незначно, що забезпечує її стабільну роботу. Завдяки цьому нейронні мережі добре підходять для задач розпізнавання образів.

Нейронні мережі вже успішно застосовуються в різних сферах: від контролю якості води та виявлення вибухових речовин у багажі до прогнозування валютних курсів у банківській сфері. Їхнє впровадження у сучасні технології відкриває нові можливості для автоматизації, аналізу великих даних і створення інтелектуальних систем.

2 РЕАЛІЗАЦІЯ КЛАСИФІКАЦІЇ ОБ’ЄКТІВ ПРИ РОЗПІЗНАВАННІ

2.1 Структура сприйняття інформації при розпізнаванні

Для усвідомленого сприйняття інформації людиною вона проходить довгий процес попередньої обробки. Спочатку світлові промені потрапляють в око, проходять через оптичну систему і досягають сітківки – шару світлочутливих клітин, що складаються з паличок і колбочок.

Ще до надходження сигналу в головний мозок відбувається перша стадія обробки: у більшості ссавців безпосередньо за світлочутливими клітинами розташовані два шари нервових клітин, які здійснюють попередній аналіз зорової інформації [9].

Далі сигнал передається зоровим нервом у мозок, зокрема в "зорові горби", де формується первинне уявлення про зображення. Потім інформація надходить до спеціалізованих відділів мозку, які виділяють основні характеристики зображення: лінії, контури, яскравість, кольори. Від простих геометричних елементів формується більш складний візуальний образ, який, перш ніж стати

частиною свідомого сприйняття, проходить ще низку трансформацій.

Аналізуючи розпізнавання образів, можна звернутися до біологічних аналогій. У певних ситуаціях можливості тварин щодо ідентифікації образів значно перевищують можливості найсучасніших комп'ютерних систем.

При розпізнаванні облич або мовлення, яке базується на сенсорному досвіді, людина значно перевершує технічні пристрої. Проте в завданнях, що потребують абстрактної логічної обробки (наприклад, класифікація за такими параметрами, як колір, форма, розмір), комп'ютерні алгоритми перевершують людей.

Оскільки процес розпізнавання образів здійснюється нейронами, можна припустити, що ключ до цього механізму слід шукати у властивостях самого нейрона. Його можна розглядати як пороговий елемент, що активується при досягненні певного порогу стимуляції.

Мак-Каллок і Піттс довели, що будь-яку обчислювану функцію можна реалізувати за допомогою організованої мережі таких нейронів. Проблема полягає в пошуку принципу самоналаштування випадково сформованої групи нейронів для виконання завдань розпізнавання. Вирішення цієї проблеми могло б стати основою теорії навчання на рівні окремого нейрона.

Канадський психолог Хебб розробив нейропсихологічну теорію навчання, яка спочатку була задумана як модель для психології, але суттєво вплинула на розвиток штучного інтелекту. Її модифіковані версії застосовувалися у створенні систем розпізнавання образів, відомих як персептрони.

Персептрони, описані Розенблаттом, можуть існувати як програмні моделі або у вигляді спеціалізованих обчислювальних пристроїв.

Образ у контексті класифікації – це група об'єктів, об'єднаних спільною ознакою. Сприйняття світу через образи є фундаментальною властивістю мозку, яка дозволяє систематизувати хаотичний потік інформації та орієнтуватися у складному навколишньому середовищі.

Людина завжди проводить класифікацію, групуючи подібні явища. Наприклад, усі написані різними почерками літери "А" або звуки однієї ноти,

зіграні на різних інструментах, сприймаються як представники однієї категорії.

Ця здатність дає змогу розпізнавати необмежену кількість варіантів після ознайомлення лише з невеликою їх частиною. Об’єктивність сприйняття підтверджується тим, що різні люди, маючи різний досвід, здебільшого однаково класифікують об’єкти. Саме завдяки цій властивості люди з різних культур можуть розуміти одне одного.

Розпізнавання образів дає змогу зіставляти сприйняті об’єкти з наявними категоріями, що робить можливим їх ідентифікацію (рис.2.1). В основі цього процесу лежить принцип узагальнення, який забезпечує впізнавання навіть варіативних зразків.

Розпізнавання образів може стосуватися як зорової (оптичної), так і звукової (акустичної) інформації. Біологічна еволюція дозволила тваринам ефективно вирішувати цю задачу за допомогою органів зору і слуху.

Створення штучних систем розпізнавання образів є складною технологічною проблемою, оскільки вимагає відтворення властивостей, притаманних природним нейронним мережам.

Проблема розпізнавання образів складається з двох ключових етапів: навчання та власне розпізнавання. На етапі навчання системі демонструються окремі об'єкти із зазначенням їхньої належності до певного класу. У результаті система повинна навчитися реагувати однаковим чином на всі об'єкти одного класу та диференціювати їх від об'єктів інших класів. Важливо, що процес навчання здійснюється на основі обмеженого набору зразків. В ролі об'єктів навчання можуть виступати зображення, літери, цифри тощо [1]. При цьому під час навчання системі повідомляється лише сам факт належності об'єкта до певного класу, без вказівки на якісь специфічні особливості або принципи класифікації.

Після навчання система переходить до процесу розпізнавання нових об'єктів, що характеризує її здатність до самостійної класифікації. Автоматизація цього процесу і становить суть проблеми навчання розпізнаванню образів.

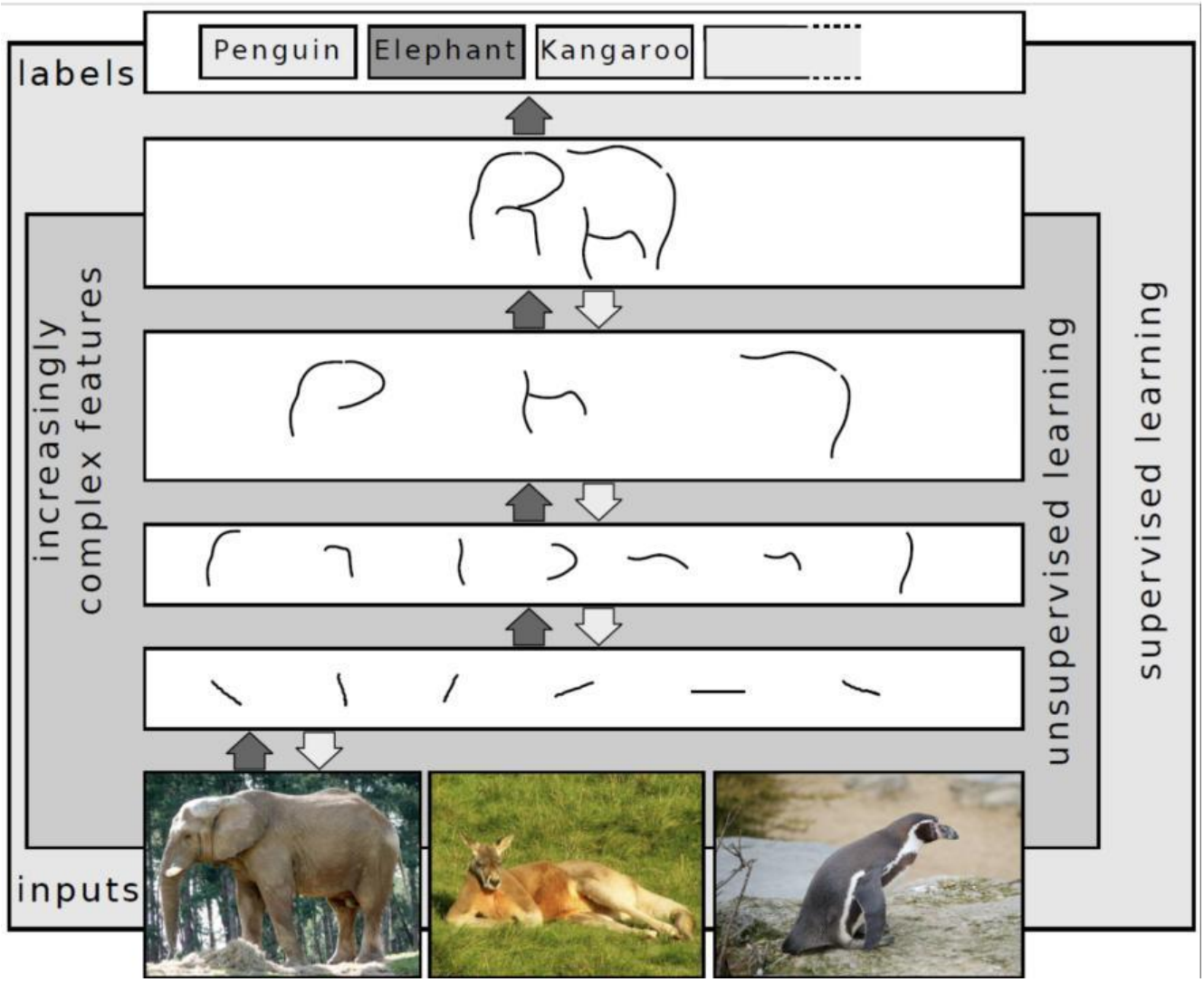


Рисунок 2.1 – Приклад об’єктів навчання

Якщо людина самостійно формулює або підбирає правило класифікації, а потім просто передає його системі, проблема вирішується лише частково, оскільки основну функцію навчання виконує людина.

Сфери застосування розпізнавання образів надзвичайно широкі. Воно використовується не тільки для ідентифікації зорових чи акустичних образів, а й у процесах класифікації складних явищ і подій. Це стосується, наприклад, прийняття управлінських рішень у бізнесі, аналізу технологічних процесів, оптимізації транспортних систем чи навіть військових стратегій. Перед початком розпізнавання необхідно отримати структуровану інформацію про об'єкт, що розглядається.

Вибір вихідного опису об'єктів є критично важливим для успішного

розпізнавання. Правильний вибір простору ознак може значно спростити задачу, тоді як невдалий опис об'єктів призведе до складного аналізу або навіть до неможливості знайти рішення. Будь-яке зображення можна подати у вигляді вектора, тобто точки в багатовимірному просторі ознак. Якщо передбачається, що всі зображення можна однозначно віднести до одного з кількох класів, це означає, що в просторі існують відповідні області, які не перетинаються.

Процес адаптації розпізнавання можна розглянути з геометричної точки зору. Задача полягає у відокремленні двох (або більше) областей у просторі. При навчанні система отримує набір точок, випадково вибраних із цих областей, та інформацію про їхню належність до певного класу. Однак жодних даних про межі областей чи правила їхнього визначення заздалегідь не надається.

Навчання має на меті побудову розділової поверхні, яка не лише чітко відокремлює відомі точки, але й забезпечує правильну класифікацію всіх інших об'єктів у цих областях. В ідеалі будується функція, значення якої є сталим для всіх точок одного класу, але різниться між класами. Оскільки області не перетинаються, завжди існує безліч можливих функцій, і процес навчання спрямований на знаходження однієї з них.

Якщо класів більше двох, завдання ускладнюється: необхідно побудувати розділову поверхню, що чітко розмежовує всі області. Це можна реалізувати за допомогою функції, яка має однакові значення в межах однієї області та відрізняється між областями.

З першого погляду здається, що знання лише обмеженого набору точок недостатньо для точного відокремлення всієї області. Існує безліч різних варіантів областей, що містять ці точки, та відповідних розділових поверхонь. Проте в математиці є класичні задачі апроксимації функцій, що дозволяють наближено визначати невідому функцію на основі її значень у кількох точках.

Важливим припущенням для розв'язання цієї проблеми є гіпотеза компактності образів. Чим компактніше та віддаленіше одна від одної розташовані області в просторі ознак, тим легше їх розмежувати. Наприклад, якщо межі областей можуть бути розділені простою площиною (рис. 2.2а), процес

розпізнавання буде більш надійним. Натомість, якщо межі складні та нелінійні (рис. 2.2б), навіть незначні помилки в їхній побудові можуть призвести до значних проблем у класифікації.

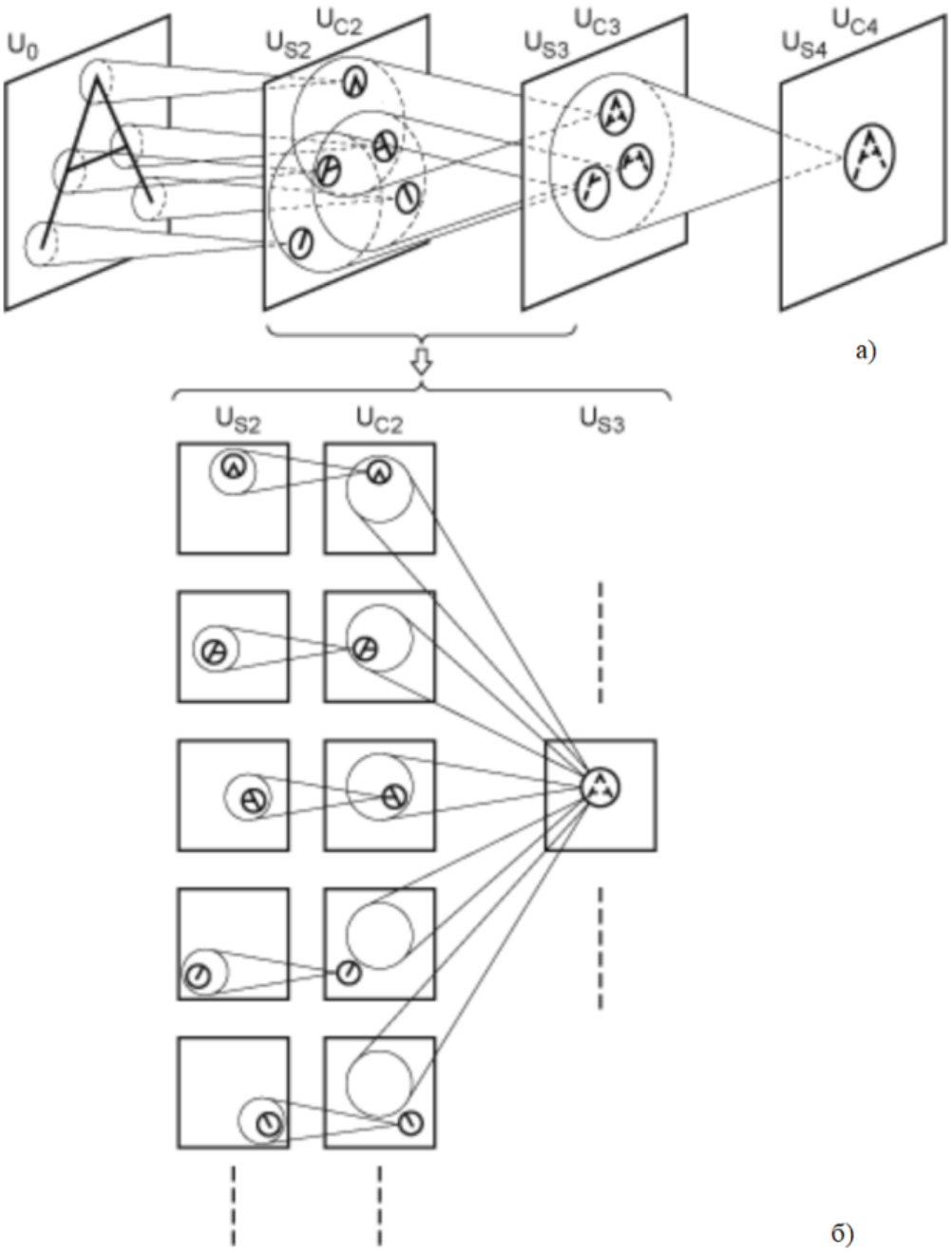


Рисунок 2.2 – Розподіл двох образів у просторі

Окрім геометричного підходу, існує ще один метод - структурний або лінгвістичний. Цей підхід широко застосовується для розпізнавання зображень.

Його суть полягає у виділенні набору базових фрагментів, які часто зустрічаються у зображеннях, а також відносних характеристик їхнього розташування (наприклад, «зліва», «вище», «усередині» тощо). Ці базові елементи формують «словник», за допомогою якого можна будувати логічні висловлювання. Завдання розпізнавання полягає у виборі з великої кількості можливих висловлювань тих, що є найбільш значущими для конкретного випадку.

Далі, переглядаючи обмежену кількість об'єктів кожного образу, необхідно побудувати їх опис. Опис має бути достатньо детальним, щоб вирішити питання про приналежність конкретного об'єкта до певного образу. У рамках лінгвістичного підходу виникають два основні завдання: створення початкового словника, що включає набір типових фрагментів, та формулювання правил для опису з елементів цього словника.

Лінгвістичний підхід передбачає аналогію між структурою зображень та синтаксисом мови. Це прагнення зумовлене можливістю використання методів математичної лінгвістики, які мають синтаксичну природу. Однак застосувати методи лінгвістики до опису зображень можна тільки після їх попередньої сегментації на складові частини. Після цього на основі цих елементів формуються "слова" для опису типових фрагментів, а завданням стає пошук цих фрагментів. Після виділення фрагментів постає завдання автоматичного граматичного розбору описів для розпізнавання зображень, що вимагає знань математичної лінгвістики та специфічних прийомів для обробки зображень.

Якщо припустити, що в процесі навчання простір ознак формується відповідно до класифікації, то завдання у цьому просторі ставить умови, при яких образи легко розділяються. Гіпотеза компактності стверджує, що компактним образам відповідають компактні множини в просторі ознак. Компактне множення – це "згустки" точок в просторі зображень, між якими є розріджені області. Хоча ця гіпотеза не завжди підтверджується експериментально, завдання, для яких вона справедлива, часто мають просте рішення. Водночас для завдань, де гіпотеза не виконується, рішення або не існує, або є складним. Цей факт спонукає до сумнівів щодо загальної правильності гіпотези компактності, однак навіть наявність

одного контрприкладу не заперечує її в тих випадках, де вона дійсно дає гарні результати. Таким чином, гіпотеза компактності залишається важливою для визначення здатності вирішувати задачі розпізнавання.

Навчання – це процес, в якому система поступово здобуває здатність відповідати на певні зовнішні впливи, а адаптація – це процес налаштування параметрів і структури системи для досягнення оптимальних результатів в умовах змінюваного середовища. У задачах навчання передбачено кілька прикладів, які служать навчальними послідовностями для правильної класифікації (рис.2.3). Якби було помічено загальну властивість, яке не залежить від природи образів, а лише від їх здатності до поділу, то можна було б сформулювати задачу навчання без учителя. Це завдання полягає в тому, що системі одночасно або по черзі показуються об'єкти без вказівок на їх приналежність до конкретного образу. Система має сама класифікувати ці об'єкти, використовуючи внутрішньо закладену властивість роздільності образів.

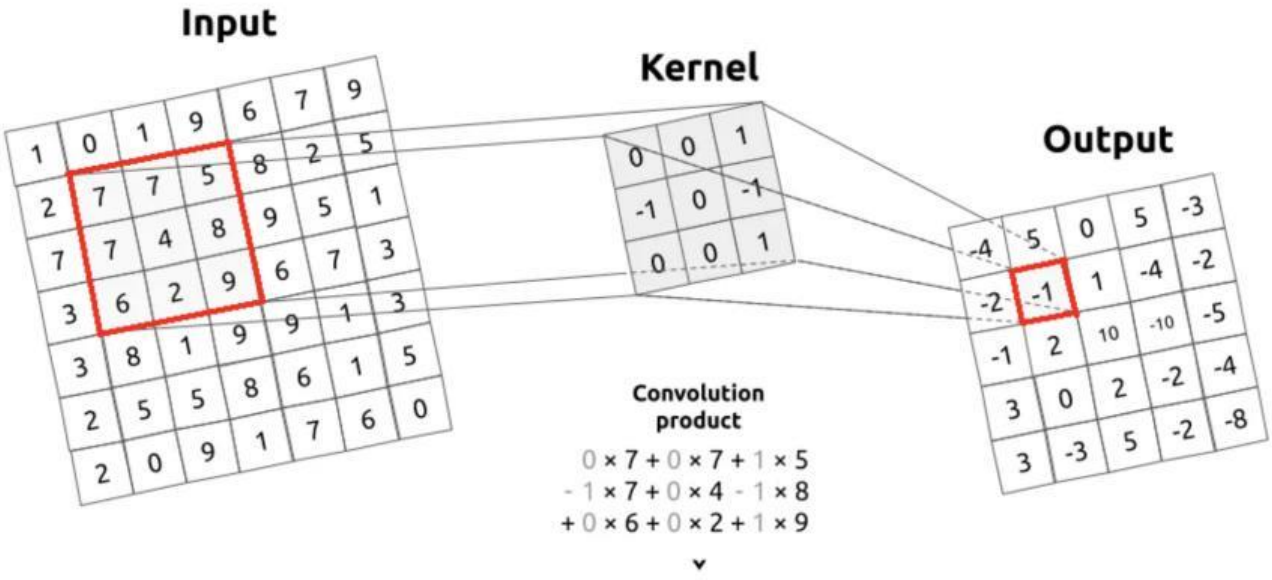


Рисунок 2.3 – Схема розпізнавання цифр

Після такого самонавчання система повинна мати можливість розпізнавати не тільки вже знайомі об'єкти (із навчальної послідовності), але й нові, раніше не пред'явлені. Самонавчання означає процес, у результаті якого система, без

допомоги вчителя, здобуває здатність правильно класифікувати об'єкти одного образу і розрізняти їх від об'єктів інших образів. У цьому процесі вчитель надає лише загальну підказку, яка допомагає системі досягти необхідної здатності до класифікації. Важливим фактором цього процесу є властивість компактності образів, оскільки саме взаємне розташування точок в просторі ознак визначає поділ цих образів.

Навчання зазвичай розглядають як процес формування в системі певної реакції на групи ідентичних зовнішніх сигналів шляхом багаторазового впливу на систему зовнішньої корекції. Цю корекцію, як правило, позначають термінами "заохочення" та "покарання". Механізм генерації цієї корекції практично повністю визначає алгоритм навчання. На відміну від навчання, самонавчання не передбачає надання системі додаткової інформації про правильність її реакцій.

Більшість відомих алгоритмів самонавчання здатні виділяти лише абстрактні образи, тобто компактні множини в заданих просторах. Відмінності між цими алгоритмами полягають у формалізації поняття компактності. Результати самонавчання свідчать про те, наскільки обраний простір підходить для вирішення конкретної задачі розпізнавання образів. Якщо абстрактні образи, що виділяються в процесі самонавчання, збігаються з реальними, то вибраний простір виявляється вдалим. Чим більше абстрактні образи відрізняються від реальних, тим менш зручним є обраний простір для конкретної задачі.

Адаптація є процесом зміни параметрів і структури системи або навіть впливів на систему з метою досягнення певного стану в умовах невизначеності та змінюваних умов роботи.

Одним із можливих способів побудови розпізнавальних машин є розрізнення ознак, властивих фігурам. Як ознаки можуть бути вибрані різні характеристики фігур, такі як геометричні (наприклад, форми кривих, з яких складається фігура) чи топологічні (наприклад, взаємне розташування елементів фігури). Відомі розпізнавальні машини, що використовують так званий "метод зондів", при якому для розпізнавання букв або цифр використовується кількість перетинів контурів фігур з кількома спеціально розташованими прямими.

Якщо проектувати цифри на поле із зондами, то виявляється, що кожна цифра перетинає певні зонди, і ці комбінації перетину зондів будуть різними для всіх десяти цифр. Ці комбінації й використовуються як ознаки для розпізнавання цифр. Такі машини успішно вирішують завдання, наприклад, для розпізнавання машинописного тексту, але їх можливості обмежуються певним шрифтом або групою подібних шрифтів, для яких була створена система ознак.

Процес створення еталонних фігур або системи ознак має здійснюватися людиною. Якість роботи машини, або надійність її "впізнавання" фігур, безпосередньо залежить від якості цієї попередньої підготовки, і без втручання людини її покращити неможливо. Така машина не є самонавчальною.

Щоб ввести зображення в машину, його потрібно перевести на машинний формат, тобто закодувати у вигляді певної комбінації символів, з якими машина може працювати. Кодування плоских фігур може здійснюватися різними способами. Найкраще прагнути до найбільш "природного" методу кодування зображень. У цьому випадку можна малювати фігури на полі, яке розбите вертикальними та горизонтальними лініями на однакові елементи - квадратики. Ті елементи, на які потрапляє зображення, будуть зафарбовані, а решта залишатимуться білими. Чорні елементи позначатимемо одиницею, а білі – нулем (рис.2.4). Потім введемо послідовну нумерацію всіх елементів поля (наприклад, у кожному рядку зліва направо, по рядках зверху вниз). Таким чином, кожна фігура, намальована на такому полі, буде однозначно представлена кодом, що складається з такої кількості цифр (0 і 1), скільки елементів містить поле.

Такий спосіб кодування вважається "природним", оскільки поділ зображення на окремі елементи відповідає принципу роботи людського зорового апарату. Сітківка ока складається з численних світлочутливих елементів - паличок і колбочок, які через нервові волокна передають інформацію у зорові центри головного мозку.

Кожен із цих елементів реагує на рівень освітленості відповідної ділянки зображення, створеного оптичною системою ока. У результаті зображення розбивається на окремі частини, що у вигляді певного коду надходять у мозок.

Такі елементи називають рецепторами, а всю сукупність - рецепторним полем.

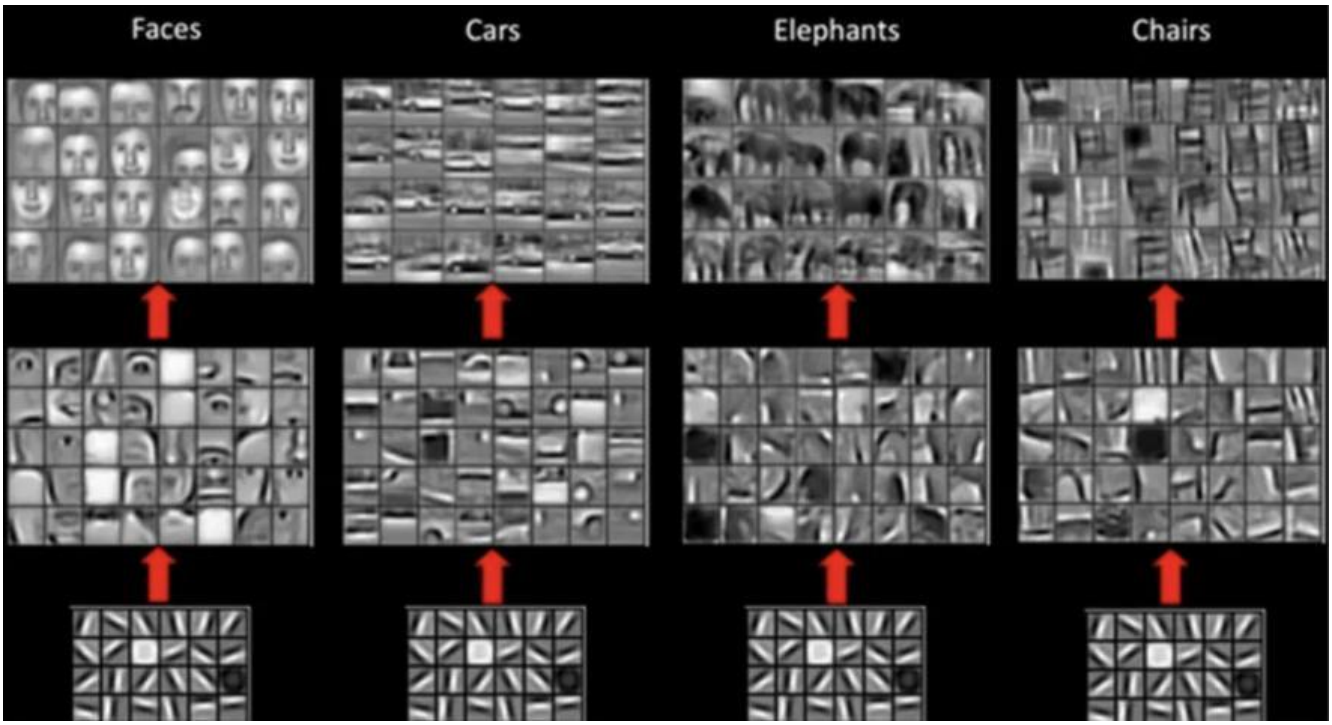


Рисунок 2.4 - Приклади нейронні мережі в процесі тренування.

Усі можливі плоскі фігури, які можуть бути відображені на рецепторному полі, утворюють множину. Кожне конкретне зображення є окремим об'єктом цієї множини. Кожному об'єкту відповідає певний унікальний код, і навпаки, кожен код можна зіставити з відповідним зображенням на рецепторному полі. Завдяки взаємно однозначній відповідності між кодами та зображеннями можна працювати виключно з кодами, зберігаючи можливість відтворення вихідного зображення.

Під ємністю нейронної мережі (ІНС) розуміється кількість образів, які вона здатна розпізнавати. Для класифікації вхідних образів, наприклад, за двома класами, достатньо одного вихідного нейрона: його значення "1" та "0" будуть позначати належність до одного з двох класів. Якщо додати ще один вихід, можна кодувати вже чотири класи, і так далі.

Щоб підвищити точність класифікації, доцільно ввести надмірність, виділивши для кожного класу окремий нейрон у вихідному шарі або, що ще ефективніше, кілька нейронів, кожен з яких визначатиме належність образу до

класу із певним рівнем достовірності (наприклад, високим, середнім чи низьким). Такі нейронні мережі здатні класифікувати образи, які належать до нечітких або перехресних множин, що наближує їхню роботу до реальних умов.

2.2 Механізм побудови нейромережі

Алгоритмічна універсальність електронно-обчислювальних машин (ЕОМ) означає, що вони здатні виконувати будь-які алгоритми обробки інформації, які можуть бути виражені у вигляді програм. Це стосується не лише обчислювальних задач, але й алгоритмів управління, пошуку та доведення теорем, генерації тривимірної графіки або створення аудіо композицій [10].

Проте не варто припускати, що комп’ютери та роботи можуть вирішувати абсолютно всі задачі. Доведено, що існують класи задач, для яких неможливо побудувати єдиний ефективний алгоритм, що вирішує всі варіанти цих задач. Відповідно, такі задачі залишаються нерозв'язними навіть для ЕОМ. Це розуміння допомагає чітко усвідомити межі можливостей машинної обробки інформації.

Однією з ключових особливостей штучних нейронних мереж (ШНМ) є їх здатність до навчання. Існує багато різних методів навчання, які часто мають схожі характеристики та засновані на спільних принципах. Процес навчання ШНМ нагадує інтелектуальний розвиток людини, хоча й має свої обмеження. Проте вже досягнуті значні результати, зокрема розробка "мовної мережі" Сейновскі, а також численні інші практичні застосування. Нейромережа навчається таким чином, щоб при певних вхідних даних формувати необхідний вихідний результат. Як вхідні, так і вихідні дані можуть бути представлені у вигляді векторів, а сам процес навчання полягає в поступовому коригуванні вагових коефіцієнтів нейронних зв’язків відповідно до заданої процедури. У міру навчання ваги мережі налаштовуються так, щоб кожен вхідний вектор давав потрібний вихідний вектор.

2.2.1 Методи навчання нейромереж

Нейромереж поділяються на два основні типи: навчання з учителем і без учителя.

Перший тип передбачає наявність "учителя", який подає вхідні образи, порівнює результуючі виходи мережі з необхідними значеннями і коригує ваги для зменшення похибки. Для кожного вхідного вектора існує відповідний цільовий вектор, що представляє бажаний вихідний результат. Разом вони утворюють навчальну пару. Під час навчання мережі подаються множини таких навчальних пар: вхідний вектор проходить через мережу, отриманий результат порівнюється з еталонним, і помилка передається назад через нейронні зв’язки, що викликає зміну вагових коефіцієнтів у напрямку мінімізації цієї похибки. Процес триває до тих пір, поки загальна помилка у всіх вхідних даних не стане прийнятно малою [11].

Навчання з учителем часто критикується через його біологічну неправдоподібність. У реальному мозку важко уявити механізм, що виконує пряму корекцію помилок за допомогою зворотного зв’язку. З цієї точки зору більш природним виглядає навчання без учителя.

Цей метод, розроблений Кохоненом та іншими вченими, не потребує заздалегідь визначених правильних відповідей. Навчальний процес використовує лише вхідні вектори, а алгоритм навчання адаптує ваги нейромережі таким чином, щоб близькі вхідні образи формували схожі вихідні відповіді. Це дозволяє мережі самостійно виділяти статистичні закономірності у навчальній множині та групувати подібні вектори в класи. Таким чином, при подачі нового вхідного вектора мережа відносить його до одного з виділених класів, хоча до навчання не можна передбачити, яким буде вихідний результат. Отже, отримані виходи потрібно додатково обробляти, щоб вони мали змістовне значення.

Процес навчання нейромережі для нового класу задач проходить через такі основні стадії:

- формулювання задачі та визначення ключових параметрів, що описують предметну область;

- вибір архітектури нейромережі, яка включає модель входів, порогову функцію, структуру шарів та алгоритми навчання. зазвичай сучасні нейромережеві програмні пакети дозволяють реалізовувати кілька базових моделей;

- підготовка навчальних даних у вигляді наборів вхідних векторів, кожен із яких пов’язаний із відповідним вихідним значенням. вхідні дані можуть бути частково неповними або суперечливими;

- навчальний процес, під час якого вхідні вектори послідовно подаються в мережу, порівнюються з еталонними значеннями, а ваги коригуються для мінімізації помилок;

- завершення навчання, коли загальна помилка досягає прийнятного рівня або мережа переходить у стаціонарний стан. один із найпоширеніших методів навчання - алгоритм "зворотного поширення помилки" (error backpropagation), що є класичним у нейроматематиці.

Після завершення навчання нейромережа може працювати з реальними вхідними даними, не лише надаючи результати, але й оцінюючи їхню достовірність.

Існує багато методів, які дозволяють машині розпізнавати образи. Один із таких методів - алгоритм, заснований на використанні січних гіперплощин. Його суть полягає в тому, що розділова поверхня між класами апроксимується наближеними гіперплощинами. Навчальний процес у цьому випадку складається з таких етапів:

- формування розділової поверхні, що включає розташування січних площин та відсікання зайвих фрагментів (рис.2. 5);

- розпізнавання нових об'єктів на основі сформованої моделі.

Метод паралельних варіантів. Ще один підхід до розпізнавання образів - метод паралельних варіантів, при якому одночасно навчається кілька незалежних машин. Під час розпізнавання кожна з них класифікує об'єкт за певним класом, причому результати можуть відрізнятися. Остаточне рішення ухвалюється шляхом "голосування" - об'єкт відноситься до того класу, який визначила

більшість машин.

Таким чином, алгоритмічна універсальність ЕОМ дозволяє реалізовувати широкий спектр інформаційних процесів, проте існують задачі, що не мають загального алгоритмічного розв’язку. Штучні нейронні мережі, завдяки своїй здатності до навчання, стають потужним інструментом обробки та класифікації даних, хоча їхня ефективність значною мірою залежить від вибору навчального алгоритму та правильного підбору навчальних даних.

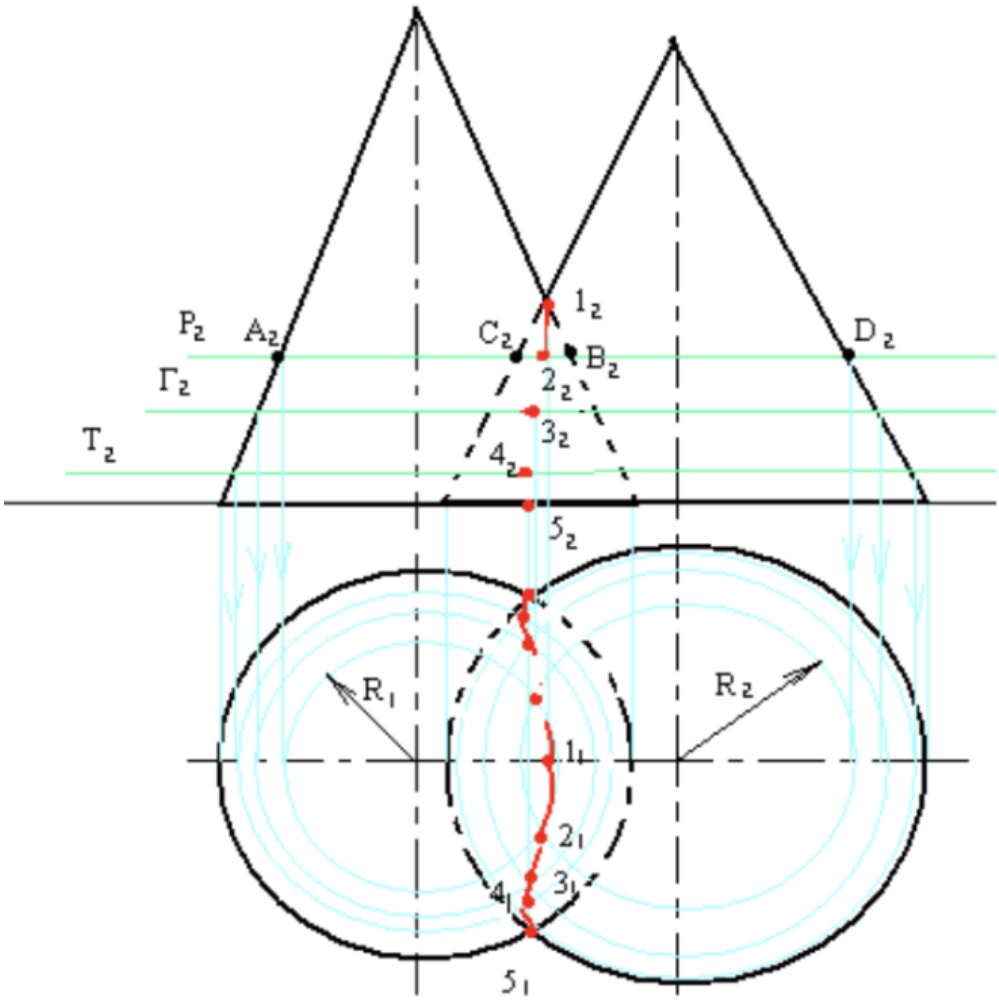


Рисунок 2.5 – Метод січних гіперплоскостей

Метод підвищення надійності розпізнавання передбачає вдосконалення процедури побудови січних площин. Припускається, що якщо розташовувати січні площини ближче до площини, яка проходить через середину відрізка, що з’єднує об’єкт із його опонентом, і є перпендикулярною до цього відрізка, то

отримана розділова поверхня буде більш точною і наближеною до істинної межі між образами. Проведені експерименти підтверджують правильність цієї гіпотези.

В алгоритмі, заснованому на методі потенціалів (рис.2.6), кожному активованому елементу рецепторного поля можна співвіднести певну функцію, яка дорівнює одиниці на цьому елементі й зменшується в усіх напрямках від нього. Ця функція φ є аналогом електричного потенціалу, з тією відмінністю, що в цьому випадку величина R визначає відстань між двома сусідніми елементами рецепторного поля.

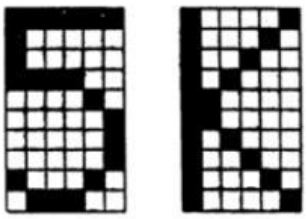


Рисунок 2.6 – Метод потенціалів

Найпростіший алгоритм розпізнавання на основі методу потенціалів реалізується у два етапи:

– навчання – у цьому процесі запам’ятовуються коди всіх точок, що з’явилися, а також вказується, до якого образу належить кожна з них;

– розпізнавання – здійснюється ідентифікація та визначається, до якого образу належить закодована матриця.

Попри певні обмеження вихідної форми персептрона Розенблатта, ця модель стала основою для створення багатьох сучасних складних алгоритмів навчання з учителем.

Прикладом персептрона з нерекурентною мережею може бути варіант, зображений на рис. 2.7. Така мережа використовує алгоритм навчання з учителем, тобто навчальна вибірка складається з набору вхідних векторів, кожен із яких має відповідний цільовий вектор. Компоненти вхідного вектора можуть набувати значень у безперервному діапазоні, тоді як компоненти цільового вектора представлені у двійковій формі (0 або 1).

Після навчання мережа отримує на вхід набір безперервних значень і видає необхідний вихід у вигляді вектора з бінарними компонентами.

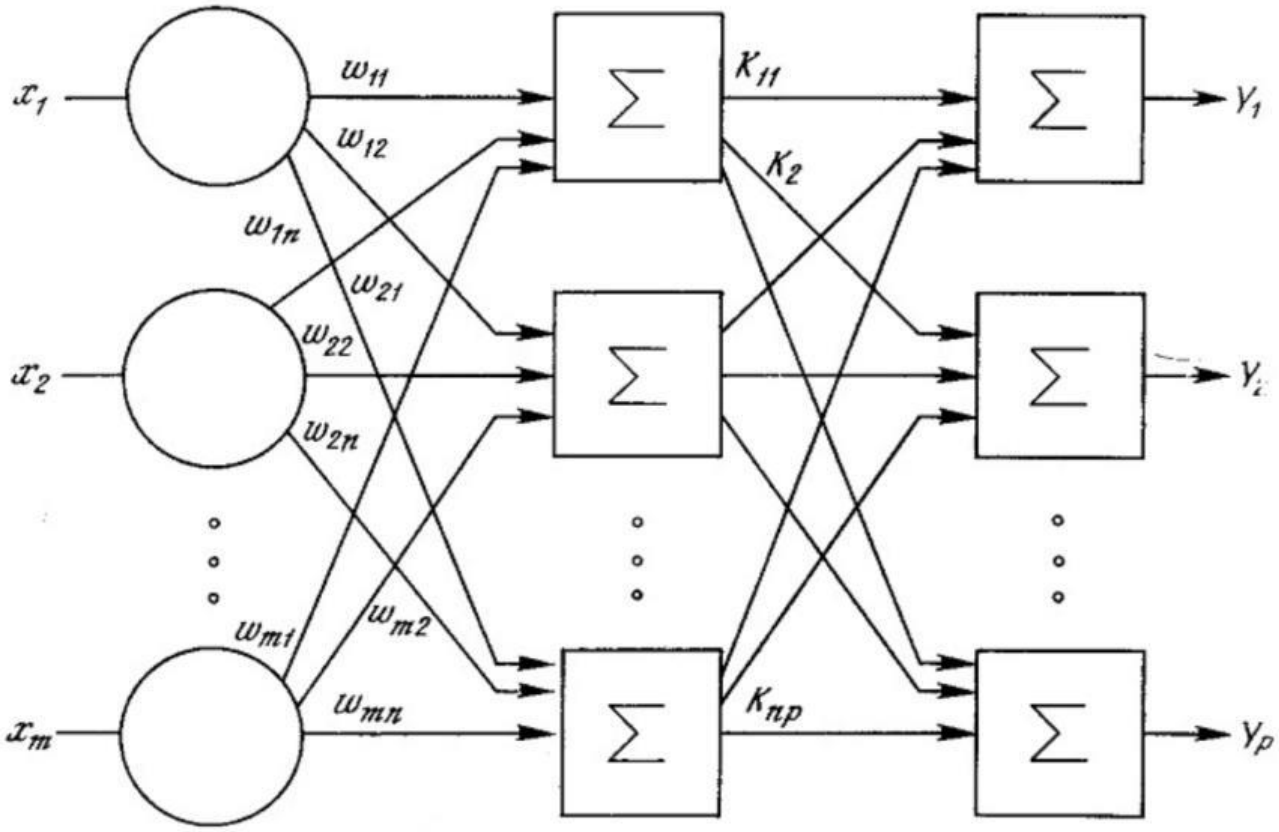


Рисунок 2.7 – Структура багатошарової нейронної мережі

Процес навчання здійснюється за такою схемою:

– ініціалізація ваг – всі вагові коефіцієнти мережі встановлюються на випадкові малі значення;

– обчислення сигналу – на вхід мережі подається навчальний вектор XX, після чого для кожного нейрона обчислюється значення сигналу NETNET за стандартною формулою 2.1;

– застосування активаційної функції – для кожного нейрона обчислюється значення порогової функції активації на основі отриманого сигналу NETNET;

– розрахунок помилки – визначається помилка кожного нейрона шляхом обчислення різниці між отриманим виходом і необхідним значенням

виходу $error_j = target_j – OUT_j$;

– оновлення ваг – кожен ваговий коефіцієнт коригується за відповідним правилом навчання, наприклад, за градієнтним методом або алгоритмом зворотного поширення помилки. Це здійснюється з урахуванням величини помилки нейрона та швидкості навчання (2.2)

– повторення процесу навчання – кроки з другого по п'ятий повторюються до тих пір, поки величина помилки не досягне прийнятно низького рівня.

$$Net_j = \sum_i x m, \tag{2.1}$$

$$W_{ij}(t+1) = w_{ij}(t) + a\, x_j\, error_j. \tag{2.2}$$

Персептрон Розенблатта має обмеження у вигляді бінарних вихідних значень. Уїдроу та Хофф розширили метод навчання персептрона на випадки, коли виходи є безперервними, використовуючи сигмоїдальну функцію активації. Вони також математично довели, що при певних умовах така мережа здатна апроксимувати будь-яку функцію, яку вона може представити. Їхня перша модель, Адалін, має один вихідний нейрон, а більш пізня модифікація – Мадалін – поширюється на випадки з множинними вихідними нейронами.

Формули, що описують навчальний процес Адалін, мають подібність до тих, що використовуються в персептронах. Однак принципова відмінність полягає в четвертому кроці: тут замість дискретних вихідних значень (0 або 1) використовуються безперервні сигнали NET. Модифікований крок 4 реалізується обчисленням помилки: для кожного нейрона визначається різниця між отриманим вихідним сигналом та необхідним вихідним значенням.

Дослідження Кохонена щодо самоорганізованих структур застосовуються для задач класифікації образів, представлених у вигляді векторних величин, де кожен компонент вектора відповідає певному елементу образу. Його алгоритми

базуються на навчанні без учителя. Після завершення навчального процесу подача вхідного вектора з певного класу призводить до активації рівня збудження у вихідних нейронах, причому нейрон із максимальним порушенням визначає класифікацію. Оскільки навчання відбувається без цільового вихідного вектора, заздалегідь неможливо передбачити, який нейрон відповідатиме певному класу вхідних векторів. Проте це можна визначити шляхом тестування мережі після завершення навчання.

У цьому алгоритмі набір з n вхідних вагових коефіцієнтів кожного нейрона інтерпретується як вектор у n-вимірному просторі. Перед початком навчання кожен компонент цього вектора ініціалізується випадковим значенням, після чого всі вагові вектори нормалізуються до довжини, рівної одиниці. Це досягається шляхом ділення кожного вагового коефіцієнта на квадратний корінь із суми квадратів усіх компонентів цього вектора.

Аналогічним чином, усі вхідні вектори навчальної вибірки також нормалізуються. Після цього мережа навчається згідно з таким алгоритмом:

– подача вхідного вектора – вхідний вектор X передається на вхід мережі;

– обчислення відстаней – визначаються відстані у n-вимірному просторі між вектором X та ваговими векторами кожного нейрона. В евклідовому просторі ці відстані розраховуються за формулою 2.3;

– визначення нейрона-переможця – нейрон, ваговий вектор якого знаходиться на мінімальній відстані від вхідного вектора XX, визначається як переможець. Цей ваговий вектор, позначений як WW, стає основним серед групи вагових векторів, що розташовані в межах відстані DD від нього;

– оновлення вагових векторів – група вагових векторів коригується відповідно до виразу 2.4;

– повторення циклу навчання – кроки з 1 по 4 повторюються для кожного вхідного вектора.

$$D_j = \sqrt{\sum_i (x - w),} \tag{2.3}$$

$$W_j(t+1) = W_j(t) + a\left[X - M\right] \tag{2.4}$$

У процесі навчання нейронної мережі параметри D і α поступово зменшуються. Початкове значення коефіцієнта α рекомендується встановлювати близько 1, з поступовим зменшенням до 0 у ході навчання. Водночас параметр DD на початковому етапі може дорівнювати максимальній відстані між ваговими векторами, а в кінцевій фазі навчання настільки зменшується, що коригується лише один нейрон.

Згідно з усталеними дослідженнями, точність класифікації покращується за рахунок додаткових навчальних циклів. За рекомендаціями Кохонена, для досягнення високої статистичної точності загальна кількість навчальних ітерацій має бути не менше ніж у 500 разів більшою за кількість вихідних нейронів.

Алгоритм навчання змінює вагові вектори в околиці активного нейрона, роблячи їх більш подібними до вхідного вектора. Оскільки всі вектори нормалізуються до одиничної довжини, вони можуть бути представлені як точки на поверхні одиничної гіперсфери. У процесі навчання група сусідніх вагових векторів наближається до вхідного вектора.

Припускається, що вхідні вектори фактично групуються в класи залежно від їхнього розташування у векторному просторі. Кожен клас асоціюється з певним нейроном, що змінює свій ваговий вектор у напрямку центру класу, збільшуючи ймовірність активації при надходженні вхідного вектора цього класу. Після завершення навчання класифікація відбувається шляхом подачі тестового вектора на вхід мережі, розрахунку рівня активації кожного нейрона та вибору нейрона з найвищим рівнем активації як індикатора правильної класифікації.

2.3 Засоби розпізнавання

Важливим етапом у розвитку нейрокібернетики став внесок нейрофізіолога Френка Розенблата, який запропонував модель розпізнавання образів, що отримала назву "Персептрон" (від латинського слова *percepto* - розумію, пізнаю). Розробляючи цю модель, він спирався на існуючі уявлення про структуру мозку та зорової системи. Намагався відтворити функції людського мозку, він використовував спрощену модель біологічного нейрона (рис. 2.8) та систему з'єднань між ними [12].

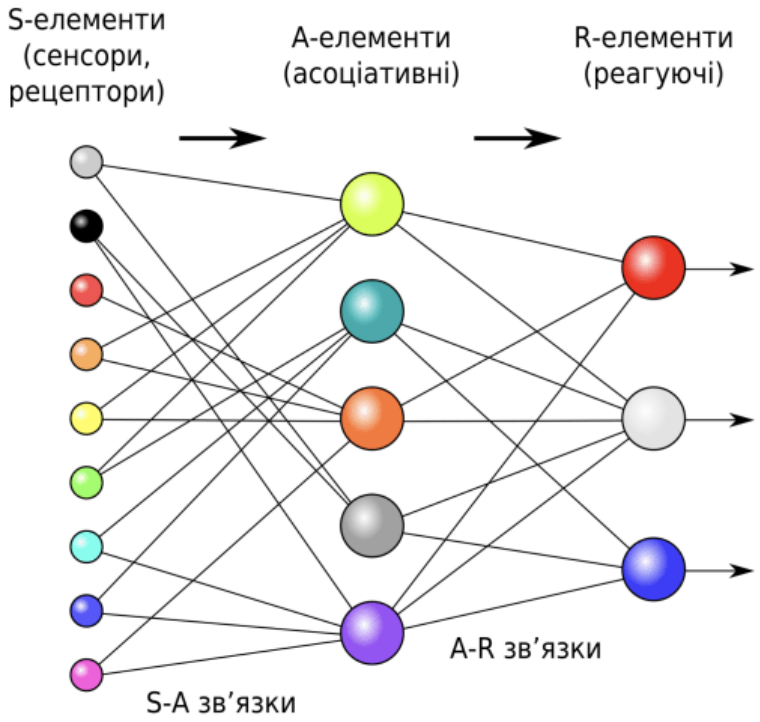


Рисунок 2.8 – Персептронний нейрон

У персептроні пристрій, що сприймає інформацію, являє собою фотоелектричну модель сітківки - поле рецепторів, яке складається з кількох сотень фотосопротівлених елементів. Кожен елемент цього поля може бути в двох станах - збудженому або незбудженому, в залежності від того, чи потрапляє на відповідний фотосопротивленій контур проекції фігури. На виході кожного

елемента генерується сигнал x_i (де i=1,2,...,n n - кількість елементів), який дорівнює одиниці, якщо елемент збуджений, і нулю в протилежному випадку. Наступним етапом у роботі персептрона є так звані асоціативні елементи або А-елементи. Кожен А-елемент має кілька входів і один вихід. Під час підготовки персептрона до експерименту виходи рецепторів підключають до входів А-елемента, при цьому кожне підключення може бути виконано з плюсом або мінусом.

Вибір рецепторів, які підключаються до конкретного А-елемента, а також вибір знака підключення здійснюється випадковим чином. Протягом експерименту зв'язки між рецепторами та А-елементами залишаються сталими. А-елементи виконують алгебраїчне підсумовування сигналів, що надходять на їхні входи, і отриману суму порівнюють з однією й тією ж величиною θ для всіх А-елементів.

Якщо сума сигналів перевищує θ, А-елемент збуджується і на його виході з'являється сигнал, рівний одиниці. Якщо сума менша за θ, А-елемент залишається незбудженим, і його вихідний сигнал дорівнює нулю. Таким чином, вихідний сигнал j-го А-елемента залежить від величини rij, яка приймає значення +1, якщо і-й рецептор підключений до входу j-го А-елемента з плюсом, значення -1, якщо з мінусом, і значення 0, якщо і-й рецептор не підключений до j-го А-елемента (j=1,2,...,m, де m - кількість А-елементів) .

Вихідні сигнали А-елементів за допомогою спеціальних пристроїв (підсилювачів) множаться на змінні коефіцієнти λj. Кожен коефіцієнт може бути позитивним, негативним або рівним нулю і змінюватися незалежно від інших коефіцієнтів. Вихідні сигнали підсилювачів підсумовуються, і отриманий сумарний сигнал надходить на вхід так званого реагуючого елемента або R-елемента. Якщо сумарний сигнал σ позитивний або дорівнює нулю, R-елемент видає на виході одиницю; якщо σ негативний - на виході з'являється нуль.

Припустимо, що на поле рецепторів проектуються фігури, що належать до двох різних класів. Якщо вдасться налаштувати персептрон так, щоб він з достатньою надійністю видавав на виході одиницю при подачі на його вхід фігур

одного класу, це буде свідчити про здатність персептрона до навчання розпізнаванню двох образів. Описана структура персептрона дозволяє класифікувати об'єкти лише на два класи. Для розпізнавання більшої кількості образів, наприклад трьох (A, B і C), можна використовувати персептрон, побудований за схемою, наведеною на малюнку. Вихідний сигнал кожного A-елемента надходить не на один, а на кілька підсилювачів, кількість яких відповідає числу різних класів. Після множення на коефіцієнти λ\lambda вихідні сигнали надходять на суматори, кількість яких також дорівнює кількості класів. Замість R-елемента використовується пристрій, що порівнює вихідні сигнали сумматорів. Об'єкт відноситься до того класу, акумулятор якого має найбільший сигнал.

Для розпізнавання кількох образів можна застосувати іншу структуру персептрона. У такій мережі A-елементи розділяються на кілька груп, кожна з яких пов'язана зі своїм сумматором і R-елементом. Сукупність вихідних сигналів від R-елементів можна розглядати як номер образу, виражений у двійковому коді, що дає змогу персептрону класифікувати об'єкти на кілька класів. Наприклад, для класифікації на вісім класів досить трьох груп. У цьому випадку можливі наступні вісім комбінацій вихідних сигналів трьох R-елементів: 000, 001, 010, 011, 100, 101, 110, 111. Кожну з цих комбінацій можна трактувати як віднесення фігури до одного з восьми класів.

Кожна з груп A-елементів, пов'язаних з певним R-елементом, функціонує так само, як і персептрон, що розділяє об'єкти на два класи. Процес навчання персептрона складається з кількох кроків. На кожному етапі персептрону подається об'єкт одного з класів, і залежно від його реакції змінюються коефіцієнти λj. За кілька етапів можна налаштувати персептрон так, що він з високою ймовірністю розпізнає подані йому фігури.

Існує два основних типи алгоритмів навчання персептрона. Перший тип не враховує правильність відповідей персептрона під час навчання, і зміна коефіцієнтів λj на кожному етапі відбувається незалежно від того, чи правильно персептрон визначив фігуру. У другому типі алгоритмів коефіцієнти змінюються

в залежності від того, чи правильно персептрон ідентифікував фігуру.

Алгоритм першого типу виглядає наступним чином: заздалегідь визначаються умови, при яких персептрон повинен видавати одиницю на виході при пред'явленні фігури класу А і нуль — при пред'явленні фігури класу В. Потім персептрону послідовно подають об'єкти з кожного класу. В кожному такті персептрон реагує на подану фігуру порушенням деяких А-елементів. Навчання полягає в тому, що коефіцієнти λj для порушених А-елементів змінюються: вони збільшуються, якщо в цей такт був поданий об'єкт класу А, і зменшуються, якщо був поданий об'єкт класу В. Така зміна коефіцієнтів сприяє покращенню правильності відповідей персептрона, оскільки збільшення λj для активованих А-елементів підвищує сигнал на вході R-елемента, а зменшення - знижує його.

Френк Розенблат, розробляючи персептрон, намагався змоделювати деякі властивості живого мозку. Персептрон, як і будь-яка програма, що імітує процес розпізнавання, працює в двох режимах: навчання та розпізнавання. На відміну від раніше розглянутих алгоритмів, алгоритм персептрона в процесі навчання не вимагає запам'ятовування об'єктів, а під час розпізнавання не перебирає всі відомі йому фігури. Це схоже на роботу мозку, який формує уявлення про образ без необхідності запам'ятовувати кожен його окремий об'єкт і здатний розпізнавати нові об'єкти без порівняння їх із усіма попередніми.

Крім того, структура персептрона має певні подібності зі структурою вищої нервової системи. Зокрема, рецептори персептрона можна порівняти з рецепторами зорового апарату, а А-елементи - з нейронами, які здатні збуджуватися, якщо інтенсивність сигналу від пов'язаних рецепторів перевищує порогову величину. Властивість персептрона мати випадкові зв'язки "рецептор - А-елемент" є аналогічною випадковому характеру зв'язків між нейронами мозку, що варіюються у різних особин одного виду. Це також дозволяє припустити, що зв'язки між нейронами не є строго фіксованими у всіх тварин цього виду, і їх варіативність може бути важливою для нормальної роботи мозку.

З точки зору стійкості, персептрон має схожість з мозком, оскільки здатний зберігати свою працездатність навіть при порушеннях його структури.

Хоча алгоритми мозку і персептрона не збігаються, персептрон є одним із найбільш правдоподібних моделей мозку на сьогоднішній день.

Здатність штучних нейронних мереж до навчання є їх ключовою характеристикою. Подібно до біологічних систем, нейронні мережі здатні самоорганізовуватися, досягаючи кращих моделей поведінки. Розенблатт створив алгоритм навчання персептрона, який забезпечує його здатність до самонавчання і дає потужний поштовх для подальших досліджень у цій галузі.

Навчання може бути як з учителем, так і без нього. У випадку навчання з учителем мережа отримує зворотний зв'язок і коригує свої ваги на основі цього зворотного зв'язку. При навчанні без учителя мережа здійснює необхідні зміни шляхом самоорганізації. Алгоритм навчання персептрона можна реалізувати на цифровому комп'ютері або іншому електронному пристрої, що робить мережу здатною до самопідстроювання. Тому процес коригування ваг зазвичай називається «навчанням», і говорять, що мережа «навчається».

3 ПРОЄКТУВАННЯ СИСТЕМИ РОЗПІЗНАВАННЯ

3.1 Порядок дій при реалізації системи

Наразі однією з важливих проблем в області інформаційних технологій є завдання розпізнавання образів з використанням ефективних методів нейрообробки інформації. Традиційні підходи до аналізу зображень та обробки даних часто не дають бажаних результатів, оскільки вони є жорстко зафіксованими та неадаптивними до змінних умов.

Існують різноманітні моделі, здатні розпізнавати друковані символи, однак їх основним недоліком є те, що вони налаштовані на конкретні образи та не володіють здатністю до самоадаптації. Тому більш універсальним рішенням було б застосування альтернатив, що дозволяють моделі самостійно налаштовуватись під нові, не визначені заздалегідь об'єкти [13].

У рамках цієї роботи поставлено завдання розпізнавання рукописних символів за допомогою комплексної моделі, що включає штучну нейронну мережу, яка поєднує поділену модель нейронів та ймовірнісно-статистичний аналіз растрових зображень. Для досягнення цієї мети була розроблена гібридна програмна модель, протестована на розпізнаванні арабських цифр, латинського та грузинського алфавітів. Однак ця модель не прив'язана до конкретних наборів символів і може бути застосована до будь-яких алфавітів та контурних зображень.

Рішення цієї проблеми залежить від багатьох факторів, і одним із них є формування достатньої бази даних (БД), що містить представників кожного класу образів. Точність і ефективність розпізнавання символів безпосередньо залежать від ретельно підібраних еталонів. Внесення в еталонну БД неповноцінних або сильно спотворених об'єктів може призвести до труднощів у класифікації та суттєвого зниження відсотка правильного розпізнавання.

Для забезпечення якісного розпізнавання необхідно мати достатню кількість еталонів, які дозволяють здійснювати аналітичні звірки та застосовувати алгоритми нейромережевої ідентифікації. Представники кожного символу в БД повинні відповідати таким вимогам: бути класифікованими, не мати сильних подібностей між собою та не містити значних деформацій усередині класу.

При розпізнаванні образів зображення передаються на матрицю рецепторів, що аналогічно процесу, коли зображення потрапляє на сітківку ока. Оброблене зображення далі надходить до мозку для подальшої ідентифікації та віднесення до певного класу символів. За аналогією, створюється модель, що має поле рецепторів, яке є прямокутним масивом, на якому можна зобразити різні конфігурації символів. Введення інформації в комп'ютер здійснюється через процес сканування символів за допомогою сканера, після чого зображення зберігається у вигляді графічного файлу. Інший спосіб - це створення зображення в графічному редакторі. Збережений файл є набором кодових знаків, які містять інформацію про структуру зображення. Для подальшої обробки ці кодові знаки необхідно перетворити в матричне подання, що складається з нулів і одиниць (0 - незаповнені, 1 - заповнені пікселі).

Сучасні програми для розпізнавання тексту використовують лексикон слів для звірки тексту, що дозволяє зменшити ймовірність помилок і штучно підвищити точність розпізнавання. Крім того, наявні програми здебільшого працюють під певні шрифти, що є ще одним обмеженням для універсальності рішень. Сучасні системи для дослідження та розпізнавання образів також використовують високороздільні матриці. Чим вище роздільна здатність (наприклад, 120 пікселів по горизонталі і 160 пікселів по вертикалі = 19 200 точок), тим точніше можна виявити закономірності та налаштувати модель розпізнавання. Для досягнення високої точності необхідно мати велику БД еталонних елементів, що включає не менше 1 000 варіантів одного символу. Створення та аналіз такої еталонної БД є трудомістким процесом, що вимагає значних часових і ресурсних витрат.

Наприклад, при наявності 1 000 варіантів для одного символу та роздільної здатності 19 200 точок (120x160), обсяг даних при 25%-ній заповненості матиме 4 800 000 точок. Якщо ж у базі є 100 різновидів символів (великі та малі літери, синоніми тощо), загальний обсяг інформації становитиме 480 000 000 точок, що є надзвичайно великим обсягом для обробки одним дослідником і потребує великих зусиль з боку команди операторів.

3.1.1 Алгоритм нейромережевої організації матриць для розпізнавання образів

Основне завдання полягає в тому, щоб на основі 40 заданих конфігурацій, збережених у БД для кожного символу, виявити їх характерні ознаки та здійснити розпізнавання символу, що подається для ідентифікації [14]. Це завдання можна розділити на кілька етапів:

– завантаження бази даних, що містить 2760 еталонних символів з файлу MS Excel у програму;

– введення зображення довільного символу для подальшого розпізнавання на матриці рецепторів (12x16) і збереження зображення у форматі .BMP (16 кольорів);

– зчитування введеного зображення з файлу для розпізнавання в

програму;

– перетворення зображення з двійкового формату в табличний масив, де 0 позначає незаповнені клітинки, а 1 – заповнені;

– обчислення конгруентної проміжної нейронної матриці 2-го рівня;

– розрахунок пасивної проміжної нейронної матриці 2-го рівня;

– виявлення характеристик симбіозних матриць та обчислення підсумкових значень на виходах;

– виявлення максимального значення серед виходів симбіозних матриць, визначення його коду і асоціювання з відповідним зображенням, яке було подано на вхід;

– коригування суматорної матриці 1-го рівня (з учителем або без вчителя) у разі високої похибки.

Алгоритм можна представити у вигляді блок-схеми. Штучна система, що самоорганізується (ІСНМС), побудована за такою схемою (рис. 3.1): на вхід нейронної мережі подається растрове зображення символу, яке можна уявити як матрицю рецепторів. Зображення подається у вигляді послідовності нулів та одиниць, де 0 означає незаповнені клітинки, а 1 - заповнені. Інформація з цієї матриці рецепторів передається на другий рівень - клас нейронних матриць, що містить представників і критерії оцінки для кожного образу з бази даних. Потім обчислюються проміжні ваги та застосовуються критерії відбору елементів, прописані в матричному вигляді на третьому рівні.

Наступним етапом є оцінка отриманих результатів через обчислення цільової функції для кожної матричної схеми з урахуванням різних критеріїв. Після того, як всі ваги будуть обчислені, на останньому етапі аналізуються отримані результати, і образ з найвищим значенням відповідає найбільш схожому образу.

У розробленій схемі передбачена самоорганізація мережі, що означає, що після визначення методу відбувається переналаштування ваг матриць критеріїв у циклі. Таким чином, схема не є статичною, а динамічно налаштовується і ефективно сходиться.

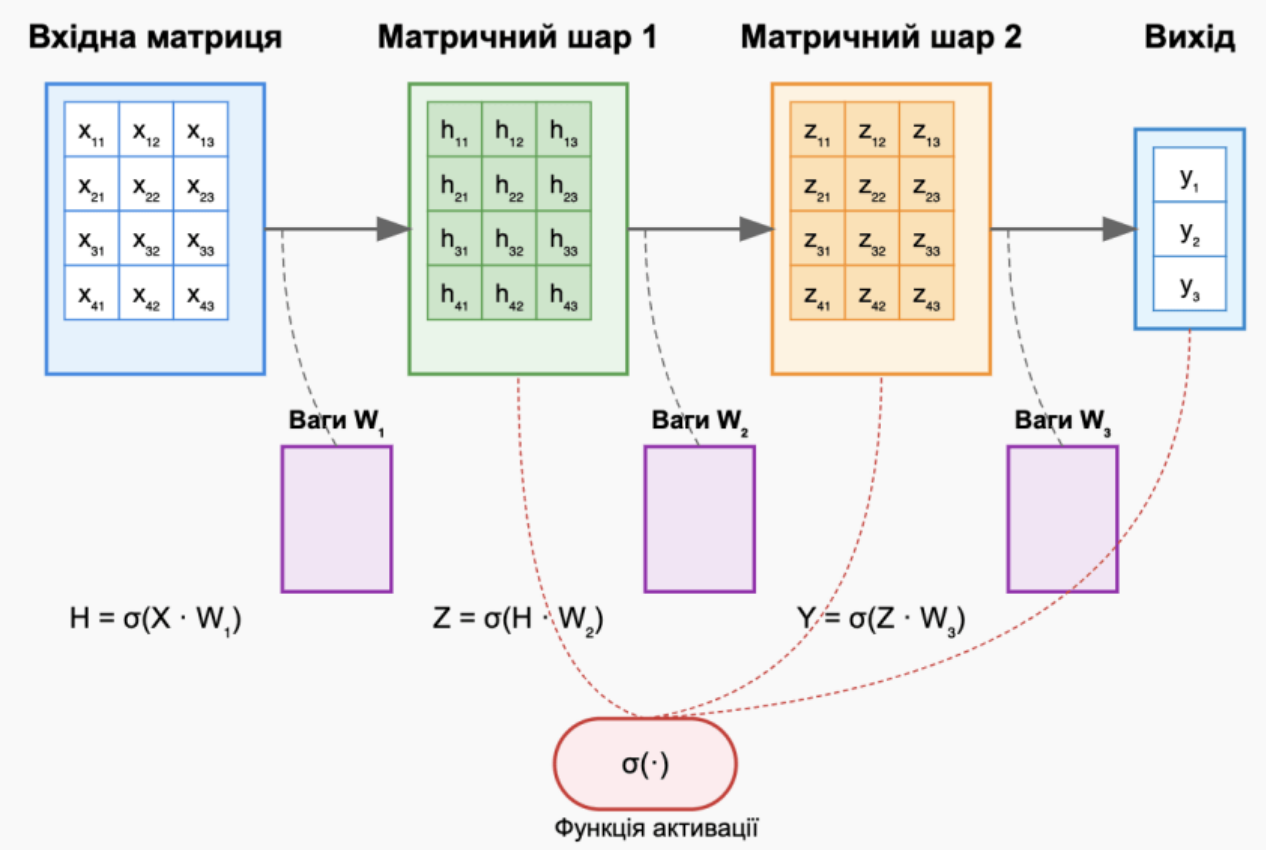


Рисунок 3.1 – Граф матричної нейронної мережі

На основі розробленої принципово нової штучної нейронної схеми з незалежними нейронами та використанням ймовірнісно-статистичних методів аналізу були досягнуті високі результати в розпізнаванні образів. Цей графік ілюструє структурну схему побудованої штучної самоорганізуючої нейронної мережі (ІСНС) для розпізнавання графічних образів.

Нижче наведена детальніша блок-схема нейронної мережі для розпізнавання образів (рис.3.2), на якій описано основні етапи, що включають завантаження бази даних, введення зображення для розпізнавання, отримання суматорних, конгруентних, діверсивних, пасивних та симбіозних матриць для кожного образу, а також обчислення цільової функції для визначення найвищого наближеного значення, що відповідає коду розпізнавання образу.

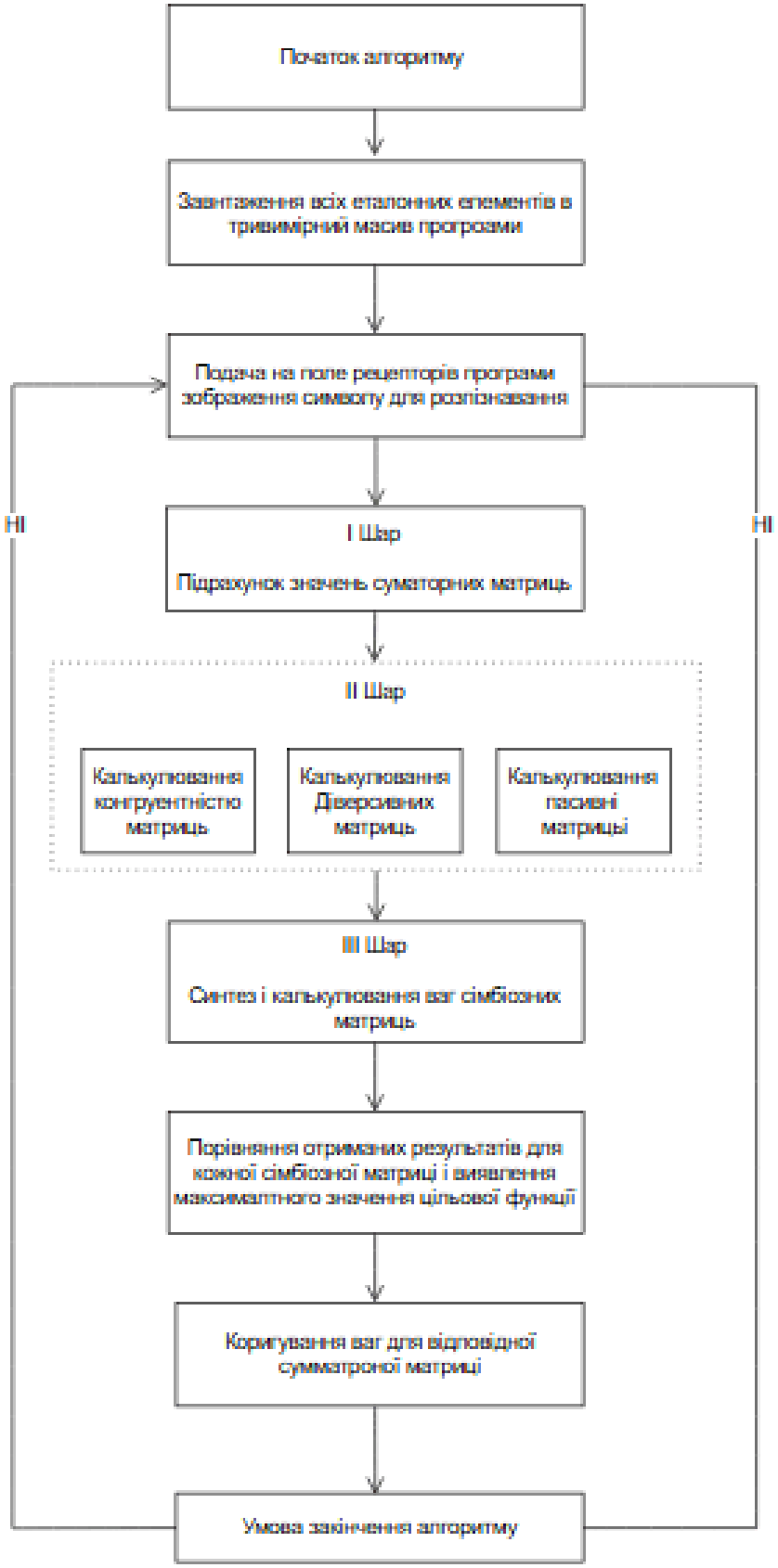


Рисунок 3.2 –Схема нейронної мережі для розпізнавання образів

3.2 Реалізація системи розпізнавання

Комп'ютерна система для розпізнавання графічних образів побудована на основі Google Cloud Vision API. Це програмне забезпечення швидко класифікує зображення на тисячі різних категорій (наприклад, "вітрильник", "лев", "Ейфелева вежа"), виявляє окремі об'єкти та особи на зображеннях, а також розпізнає і читає надруковані слова, що містяться в зображеннях (рис. 3.3). Додатково, система дозволяє створювати метадані для зображень із помірним образливим змістом або розробляти нові маркетингові стратегії за допомогою аналізу зображень. Зображення можуть бути завантажені через запит або інтегровані із сховищем зображень на Google Cloud Storage.

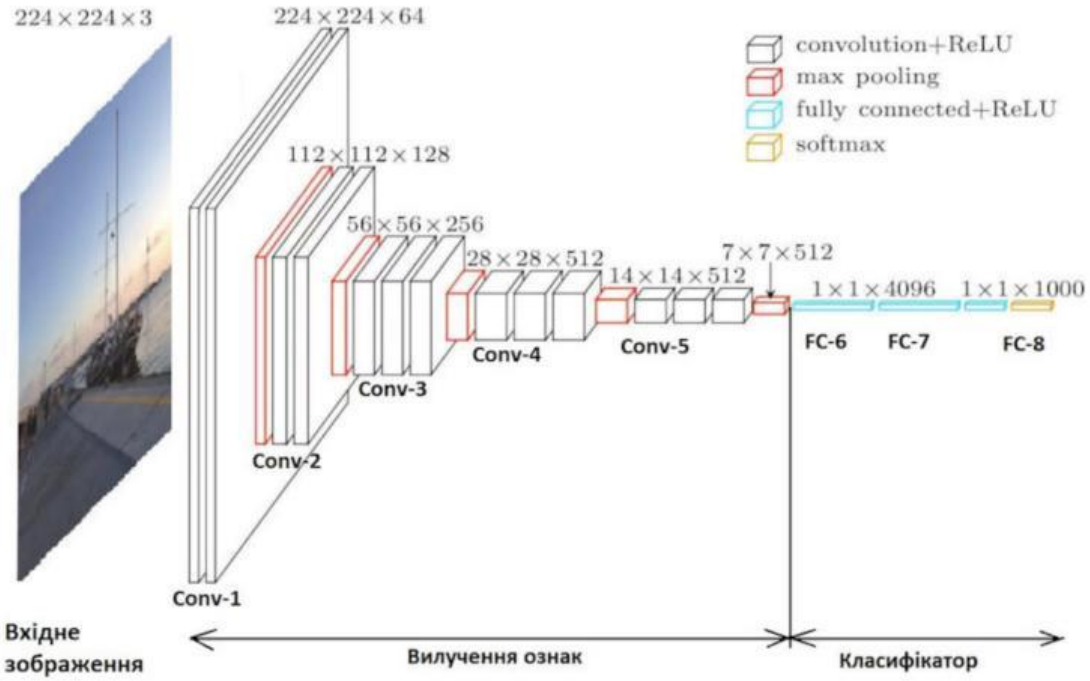


Рисунок 3.3 - Класифікація графічних образів

Ця комп'ютерна система дозволяє швидко виявляти об'єкти на зображеннях, такі як квіти, тварини або безліч інших категорій, що часто зустрічаються в зображеннях (рис.3.4). З часом Vision API додає нові концепції, що сприяє підвищенню точності розпізнавання.

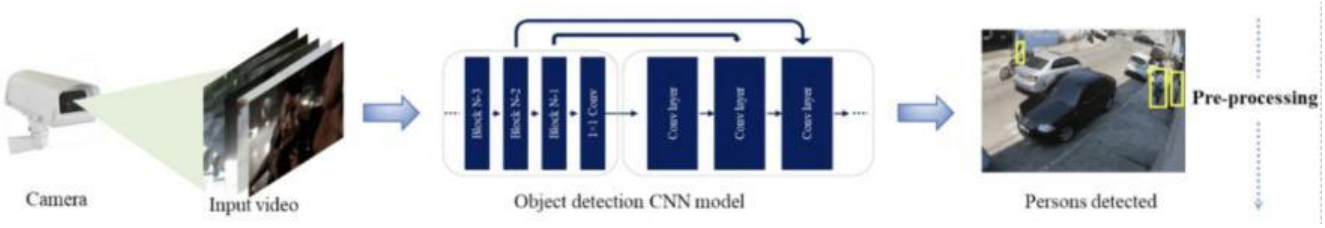


Рисунок 3.4 – Виявлення графічних об‟єктів

Розроблена комп'ютерна система складається з таких модулів [15]:

– LabelDetection – розпізнає широкий спектр категорій об'єктів на зображенні, від транспортних засобів до тварин.

– ExplicitContentDetection – виявляє наявність явного або небажаного контенту на зображенні.

– LogoDetection – ідентифікує популярні логотипи продуктів у межах зображення.

– LandmarkDetection – визначає відомі природні та штучні об’єкти, такі як пам’ятки архітектури.

– OpticalCharacterRecognition – розпізнає та витягує текст із зображень, підтримуючи широкий набір мов та автоматичне визначення мови.

– FaceDetection – виявляє обличчя на зображенні та визначає ключові характеристики, зокрема емоційний стан чи наявність головного убору. При цьому функція ідентифікації осіб не підтримується.

– ImageAttributes – аналізує загальні характеристики зображення, наприклад, домінуючий колір.

Для використання Cloud API необхідно налаштувати облікові дані програмного продукту (рис.3.5) для автентифікації та отримання дозволу на виконання операцій. Під час доступу до API Google Cloud Platform рекомендується створити ключ API для тестування та налаштувати обліковий запис служби для роботи з продуктом.

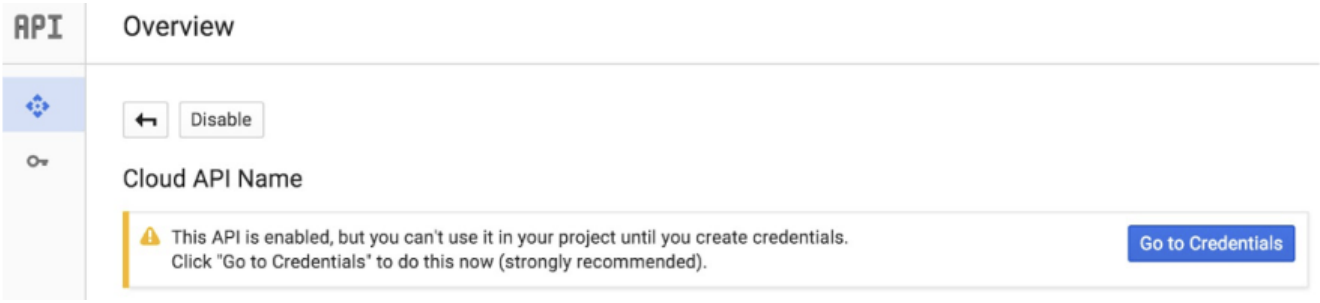


Рисунок 3.5 - Налаштування проекту

Щоб додати облікові дані, необхідно натиснути кнопку "Go to Credentials" (рис.3.6):

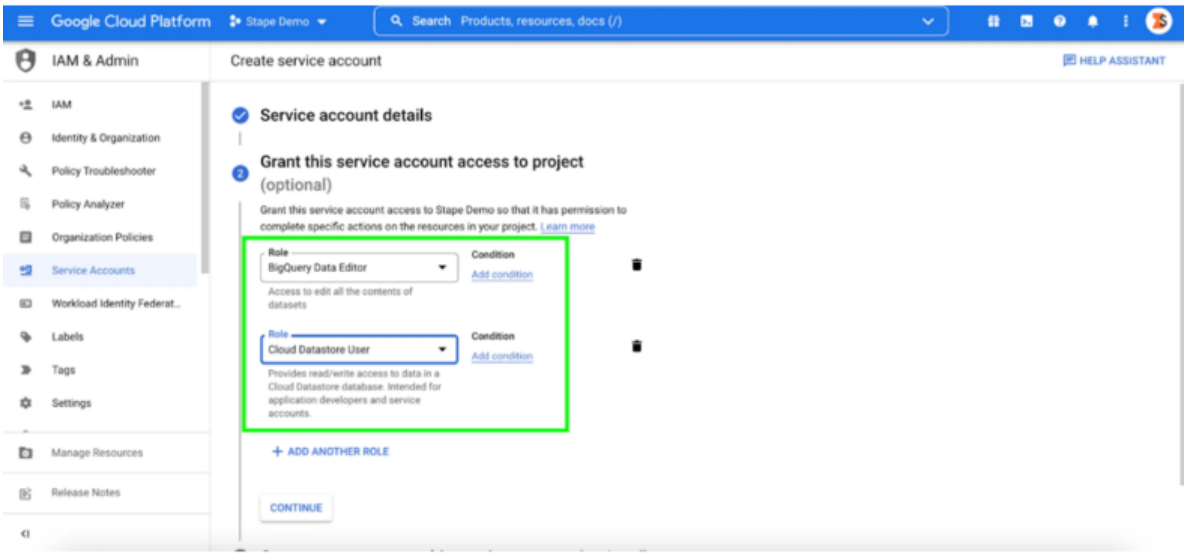


Рисунок 3.6 - Додавання даних

Аутентифікація та авторизація в Google Cloud Platform API (зазвичай об'єднуються під терміном "авт") зазвичай здійснюється через обліковий запис служби. Такий обліковий запис дозволяє додатку безпосередньо передавати дані до Cloud API. Він функціонує подібно до облікового запису користувача та представлений у вигляді електронної адреси. Однак, на відміну від звичайного користувацького акаунта, обліковий запис служби призначений виключно для додатка та використовується для доступу лише до тих API, для яких він був створений (рис. 3.7). Далі наведено приклад створення облікового запису служби за допомогою консолі Google Cloud Platform.

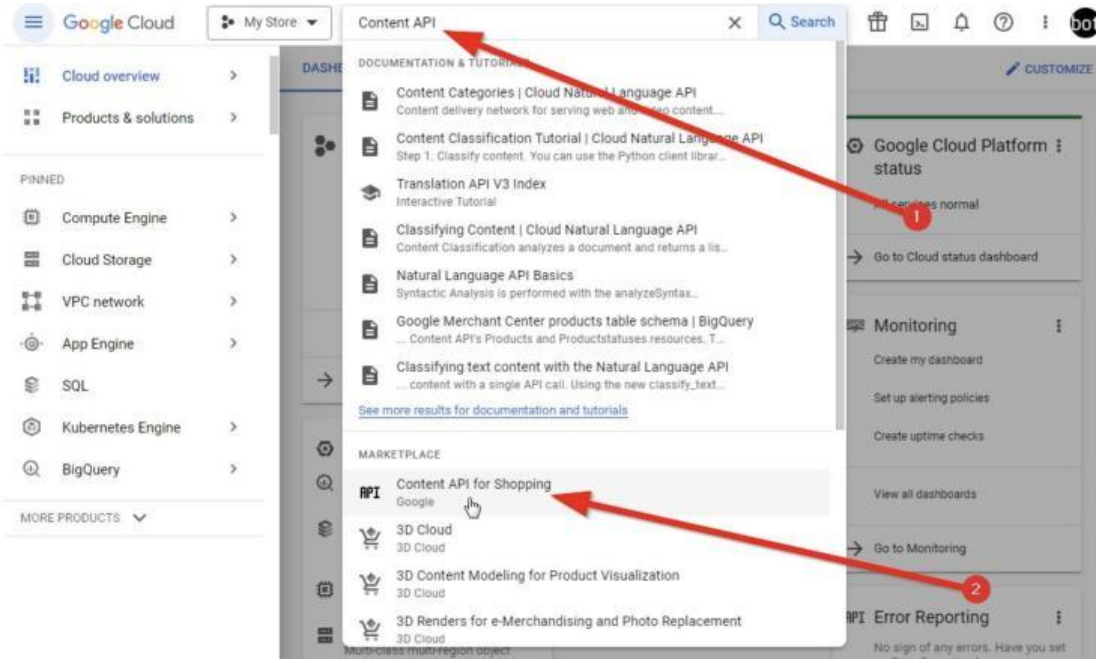


Рисунок 3.7 – Під’єднання API ключа

Розглянемо реалізацію модулів комп’ютерної системи. Модуль LabelDetection дозволяє виявляти широкий спектр категорій об’єктів на зображенні — від транспортних засобів до тварин. Нижче представлений приклад процесу розпізнавання міток у локальному файлі зображення. Звернення до images:annotate здійснюється через спеціальний API-ендпоінт. Для виконання виявлення міток необхідно надіслати POST-запит до відповідного сервісу, передавши необхідні дані.

Лістинг 3.1 - POST-запит

```
POST
https://vision.googleapis.com/v1/images:annotate?key=YOUR_API_KEY

{
  "requests": [
    {
      "images": {

        "content": "/9j/7QBEUGhvdG9zaG9...base64-encoded-image-
content...fXNWzvDEeYxxxzj/Coa6Bax//Z"
    },
    "features": [
      {
        "type": "LABEL_DETECTION"
      }
```

```
    ]
   }
  ]
}
```

Реалізація LabelDetection за допомогою бібліотеки має наступний вигляд:

Лістинг 3.2 - Реалізація LabelDetection

```
useGoogle\Cloud\Vision\VisionClient;

// $projectId = 'YOUR_PROJECT_ID';
// $path = 'path/to/your/image.jpg'


$vision =newVisionClient([
     'projectId'=> $projectId,
]);

$image = $vision-
>image(file_get_contents($path),['LABEL_DETECTION']);
$result = $vision->annotate($image);

print("LABELS:\n");
  foreach($result->labels()as $label){
     print($label->description(). PHP_EOL);
  }
```

Розглянемо приклад реалізації FaceDetection**.** Для доступу до API Google через офіційний клієнт SDK, необхідно створити об’єкт сервісу на основі документації API, яка описує його інтеграцію в SDK. Отримати цей об’єкт можна за допомогою Vision API, використовуючи відповідні облікові дані.

Лістинг 3.3 - Реалізації FaceDetection

```
useGoogle\Cloud\Vision\VisionClient;


// $projectId = 'YOUR_PROJECT_ID';
// $path = 'path/to/your/image.jpg'


$vision =newVisionClient([
     'projectId'=> $projectId,
]);
```

Щоб сформувати запит до Vision API, спочатку слід ознайомитися з його документацією. У даному випадку необхідно звернутися до ресурсу зображень**,** щоб додати коментар до зображення. Запит до API має вигляд об’єкта, що містить список запитів. Кожен елемент цього списку включає два основних компоненти.

Лістинг 3.4 - Запит до API

```
$image = $vision-
>image(file_get_contents($path),['FACE_DETECTION']);
$result = $vision->annotate($image);
```

У відповідь на запит про анотацію облич, API надає широкий набір метаданих щодо виявлених облич. Ці дані включають координати багатокутника, що окреслює обличчя. Однак на цьому етапі вони представлені лише у вигляді списку числових значень. Далі можна використати ці координати для перевірки того, що обличчя дійсно було знайдено на зображенні.

3.3 Аналіз результатів

Комп'ютерна система розпізнавання графічних образів дає змогу ідентифікувати різні типи об'єктів на зображеннях за допомогою нейронних мереж. Реалізовано функції розпізнавання міток, тексту, кольорів, облич тощо.

Розглянемо процес розпізнавання об'єктів на наведеному графічному зображенні (рис.3. 8):

Рисунок 3.8 - Базове зображення для тестування

Результати виконання даної комп’ютерної системи наведено на рисунку 3.9 та таблиці 3.1.

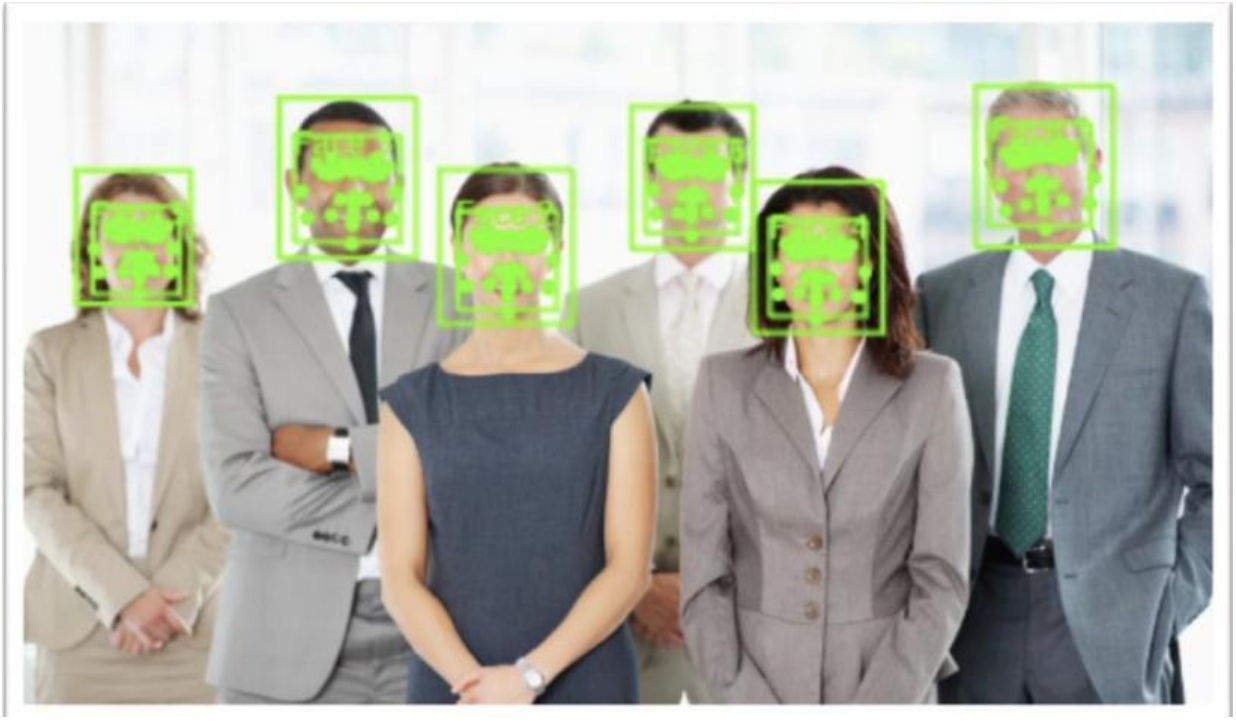


Рисунок 3.9 - Результат розпізнавання облич

Таблиця 3.1 - Результат розпізнавання емоцій на двох обличчях

Базове зображення			Результат розпізнавання					
Особа 1	★	малоймовірно	★	★	★	★	★	ймовірно
Особа 2	★	малоймовірно	★					малоймовірно
Особа 3	★	малоймовірно	★					малоймовірно
Особа 4	★	малоймовірно	★					малоймовірно
Особа 5	★	малоймовірно	★					малоймовірно
Особа 6	★	малоймовірно	★					малоймовірно
Особа 7	★	малоймовірно	★					малоймовірно
Конфіденційність		98%	Конфіденційність					100%

Дане графічне зображення було класифіковане модулем LabelDetection із міткою “Team” (рис.3.11), що є найбільш очевидним результатом під час аналізу цього зображення.

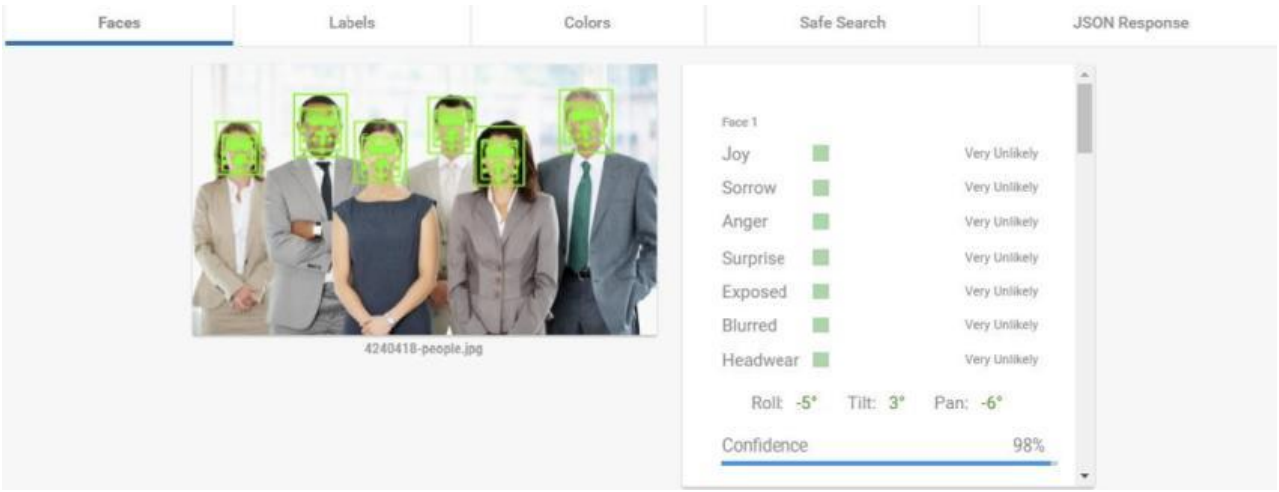


Рисунок 3.11 – Інтерфейс системи

Отже, на основі отриманих результатів можна зробити висновок, що дана система функціонує повноцінно та відповідає всім поставленим завданням. Крім того, було значно підвищено швидкодію розпізнавання графічних образів та знижено апаратні витрати.

ВИСНОВКИ

На основі проведених досліджень розроблений принципово новий метод розпізнавання образів, який використовує унікальний алгоритм та програмну реалізацію. Цей метод забезпечує функціонування нової розділеної структури штучної нейронної мережі, що базується на нейро-матричній реалізації (конгруентність, діверсивні, пасивні та нейтральні масиви растрових крапок), з використанням ймовірнісно-статистичного аналізу для класифікації оброблюваних зображень за спеціальними мітками.

Основні результати дипломного проєкту:

- проведений аналіз літератури та ресурсів Інтернету дозволив виявити й застосувати принципи з різних галузей, таких як нейробіологія, нейроінформатика, кібернетика, штучний інтелект, нейронні мережі, самоорганізація мереж, персептрони, кластеризація та категоризація об'єктів, аналіз сцен, машинне розпізнавання образів, теорія множин, дискретна математика, кодування/декодування, графічні формати файлів, теорія ймовірності та математична статистика, а також алгоритмічні побудови блок-схем та об'єктно-орієнтоване програмування на мові PHP;

- створена принципово нова штучна нейро-матрична мережа без вчителя, яка обробляє і модифікує еталонну базу образів (суматорну матрицю) при порівнянні з тестовими зображеннями. Підстроювання мережі здійснюється посимвольно, змінюючи вагові коефіцієнти двовимірних масивів з БД, з використанням нових методів адаптації для підвищення ефективності розпізнавання образів. Запроваджено нові поняття діверсивних, конгруентних і нейтральних мас. Основою методу є ймовірносно-статистичні оцінки для вагових коефіцієнтів в багатошаровій мережі, а також нова ідея змішування піксельних мас для виявлення приналежності до класифікованих груп;

- результати експериментальних досліджень, що продемонстрували досягнення високої точності класифікації при мінімальному дозволі матриці

рецепторів та кількості представників образів у базі даних, підтверджують ефективність нового методу вагових параметрів синаптичних масивів для навчання нейронної мережі. Ці результати мають практичну цінність і можуть бути впроваджені в різних галузях науки і практики;

- змодельовано, протестовано та налагоджено комп'ютерну програму на основі принципово нової комплексної структурної моделі нейронної мережі, яка використовує ймовірностатистичний аналіз нейро-матричних масивів графічної інформації з багатороздільною мережею нейронів для паралельної класифікації образів.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Басюк Т., Пушко Я. Аналіз та класифікація основних методів розпізнавання образів на площині проекції. Національний університет “Львівська політехніка”, кафедра інформаційних систем та мереж. Львів, 2015. С. 291-298.

2. How DLP Technology Works. URL: http://www.dlp.com/ technology/how-dlp-works/default.aspx (last accessed: 15.04.2025).

3. LCoS Projectors. URL: http://www.projectorreviews.com/projector categories/lcos-projectors (last accessed: 09.04.2025).

4. Bishop C. M. Neural Networks for Pattern Recognition. Oxford University Press, 1996. 504 p.

5. Довбиш А. С. Основи теорії розпізнавання образів: навчальний посібник. Суми : Сумський державний університет, 2015. Ч. 1. С. 109.

6. Б. М. Самусько В. П. Майданюк. Класифікація методів для розпізнавання зображень. URL:http://efaidnbmnnnibpcajpcglclefindmkaj/https:// ir.lib.vntu.edu.ua/bitstream/handl (дата звернення 28.04.2025).

7. Розпізнавання образів. URL: http://patterns-recognition.blogspot.com/ (дата звернення 19.04.2025).

8. Кутковецький В. Я. Розпізнавання образів: навчальний посібник. Миколаїв: Вид-во ЧНУ ім. Петра Могили, 2017. С. 420.

9. Гавриленко В.В. Теорія розпізнавання образів: навчальний посібник. К.: НТУ, 2013. С. 96.

10. Литвин В. В., Нікольський Ю.В., Пасічник В.В. Методи та засоби інженерії даних та знань. Львів: Магнолія. 2021. С. 252.

11. Object Management Group® Standards Development Organization. URL: https://www.omg.org/index.htm/ (last accessed 25.05.2025).

12 Довбиш А. С. Основи теорії розпізнавання образів: навчальний посібник. Суми : Сумський державний університет, 2015. Ч. 2. С. 132.

13. Копча-Горячкіна Г. Е. Теорія розпізнавання образів: навчально-

методичний посібник. Ужгород: Видавництво ДВНЗ «Ужгородського національного університету». 2016 р. Частина І. С.99.

14. Clustering. URL: https://scikit-learn.org/stable/modules/clustering.html/ (Last accessed: 29.04.2025).

15. Parker J. Principles of scientific research. 7th ed. London: Editorial, 2017. P.301.

Національний університет «Запорізька політехніка»
Кафедра комп'ютерних систем та мереж

Дипломний проєкт

на тему «РОЗРОБКА КОМП'ЮТЕРНОЇ СИСТЕМИ РОЗПІЗНАВАННЯ ГРАФІЧНИХ ЗОБРАЖЕНЬ НА БАЗІ НЕЙРОННИХ МЕРЕЖ»

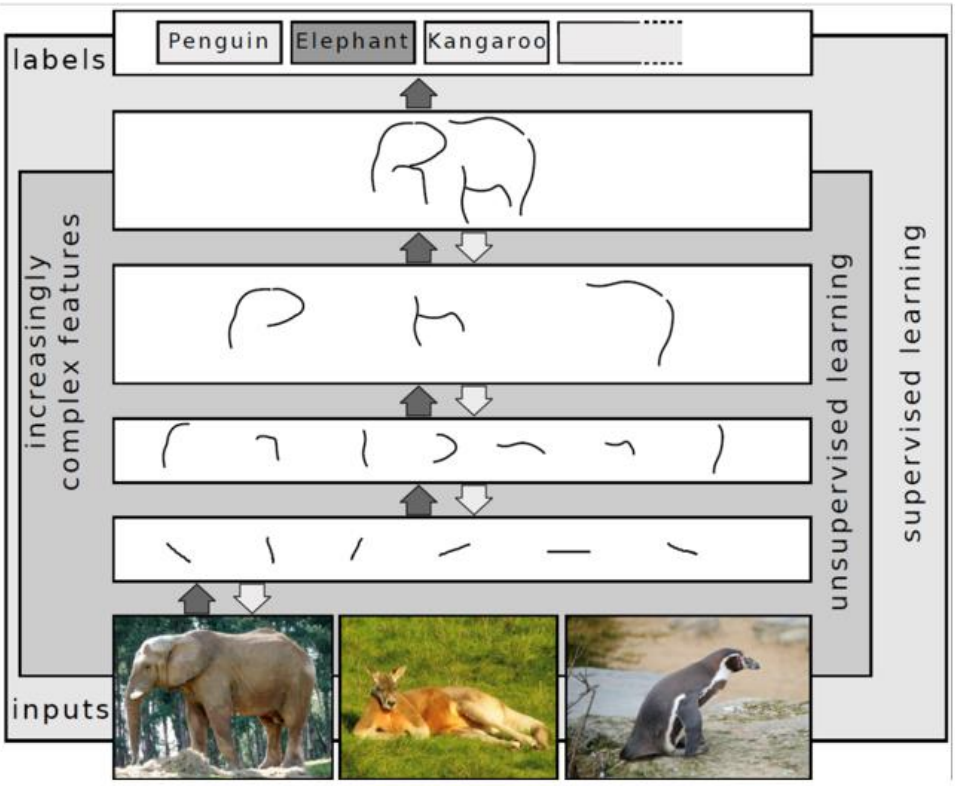
студента групи КНТ-512сп НІКІТЮК А.І
керівник проєкту ІЛЬЯШЕНКО М.Б.

МЕТА РОБОТИ - розробка, проєктування та тестування вдосконаленої нейронної мережі для спрощення процесу розпізнавання образів.

Завдання до роботи - зробити систему розпізнавання рукописних символів за допомогою комплексної моделі, що включає штучну нейронну мережу, яка поєднує поділену модель нейронів та ймовірнісно-статистичний аналіз растрових зображень.

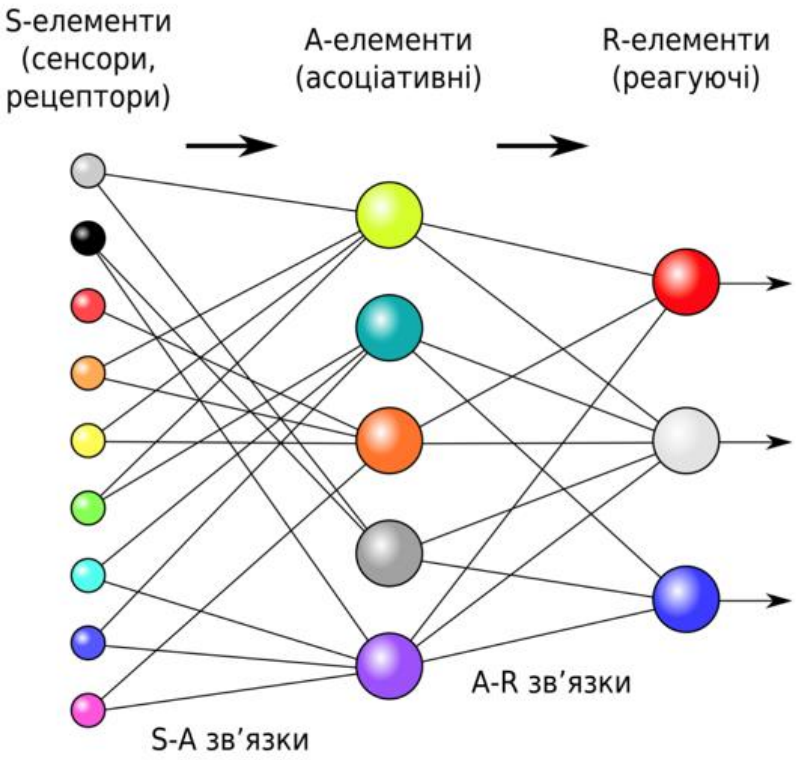
Огляд існуючих нейронних мереж і їх застосування :

- ✓ метод бінаризація зображення - процес переведення кольорового чи напівтонового зображення документа в двоколірне (бінарне, чорно-біле);
- ✓ метод фільтрації Фур'є, низькочастотні фільтри (ФНЧ) та високочастотні фільтри (ФВЧ) - використовує спектральні характеристики для виділення частот у сигналі.;
- ✓ фільтр Калмана (для низьких частот) - який використовується для зменшення впливу високочастотних шуму та коливань у сигналах, залишаючи низькочастотні складові, які можуть бути більш важливою інформацією;
- ✓ фільтр Габора (для високих частот) - це представлення перетворень зображення за допомогою фільтрів, які застосовують різні варіації масштабів і орієнтацій.

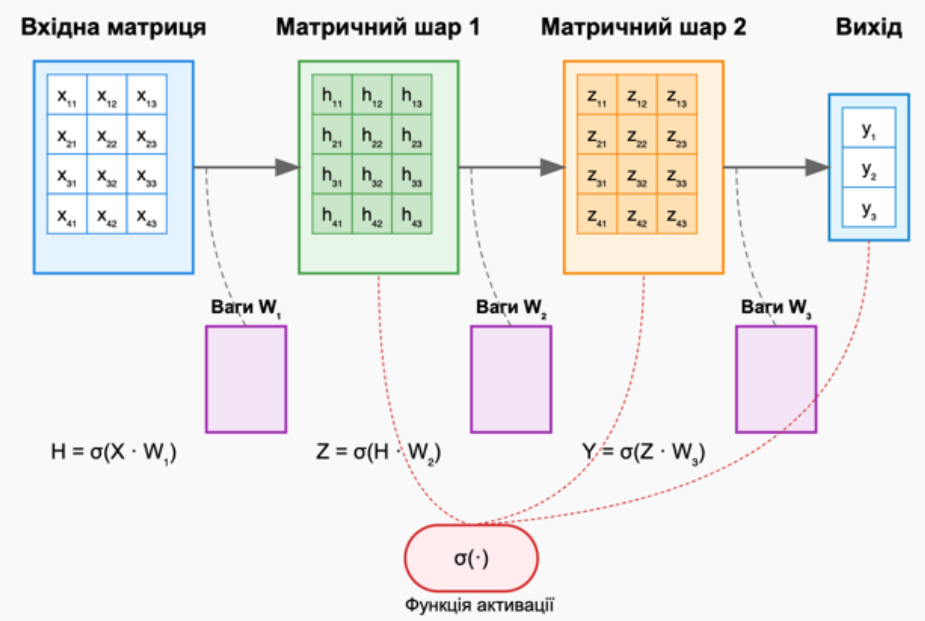


Приклад об'єктів навчання

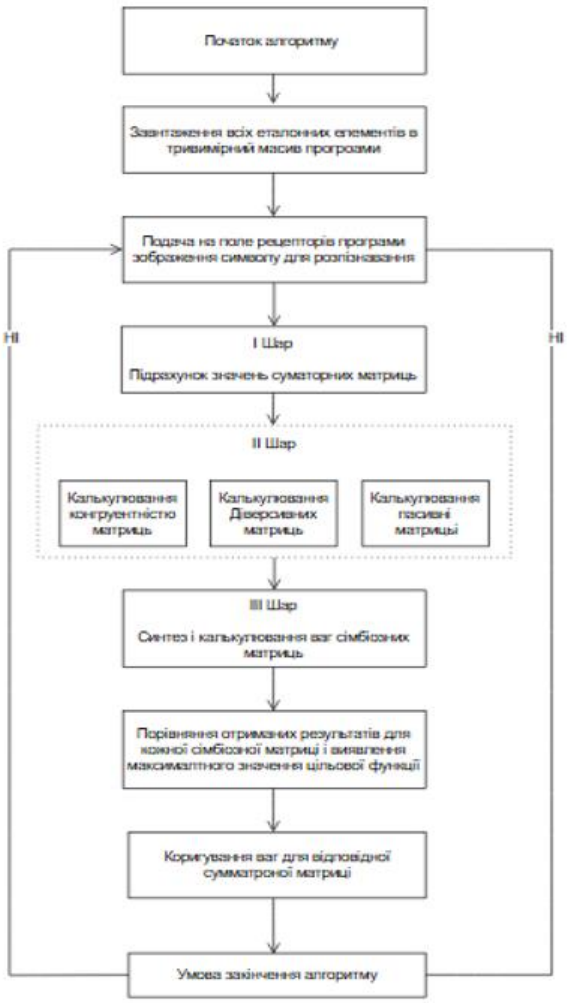
Навчання нейромережі

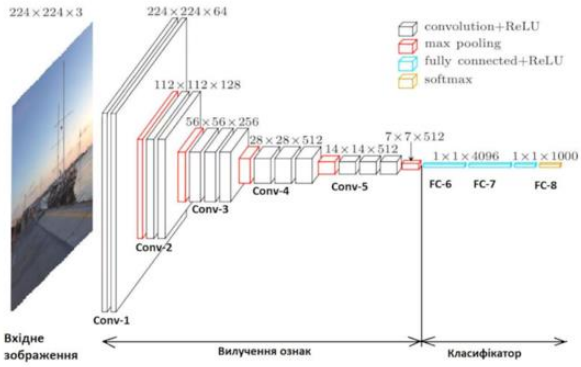


Персептронний нейрон

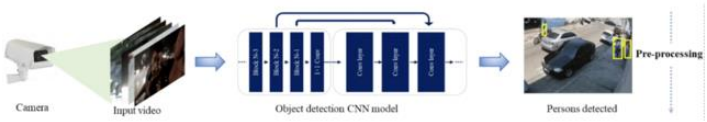


Граф матричної нейронної мережі





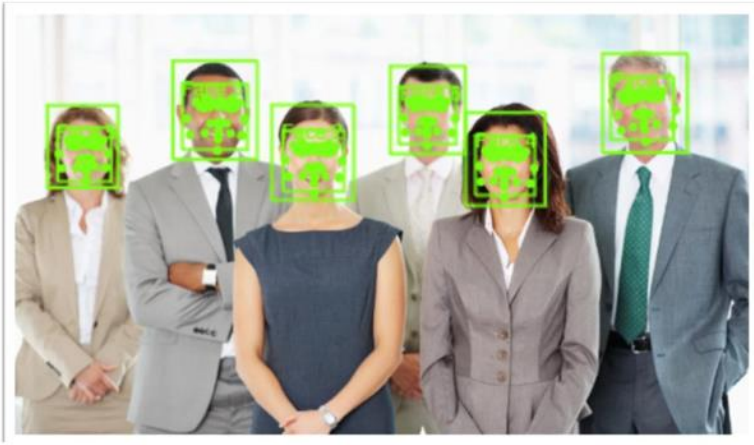
Класифікація графічних образів



Виявлення графічних об'єктів

Зображення для тестування

Результат розпізнавання облич



ВИСНОВОК

По результатам виконання дипломного проєкту було розроблено принципово новий метод розпізнавання образів, який використовує унікальний алгоритм та програмну реалізацію.

Протестовано та налагоджено комп'ютерну програму на основі структурної моделі нейронної мережі для паралельної класифікації образів.