МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
УМАНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ САДІВНИЦТВА
Факультет економіки і підприємництва
Кафедра інформаційних технологій

Допущено до захисту
"_____" _____ 2024 р.
Зав. кафедри, к.т.н., доцент
_____ **Ліщук Р.І.**

КВАЛІФІКАЦІЙНА РОБОТА МАГІСТРА

РОЗРОБКА ІНФОРМАЦІЙНОЇ СИСТЕМИ ДЛЯ РОЗПІЗНАВАННЯ ТЕКСТОВИХ ЗОБРАЖЕНЬ ІЗ ЗАВАДАМИ

Виконав – студент 2 курсу, групи 21м-кн
освітньої програми "Комп’ютерні науки"
спеціальності 122 Комп’ютерні науки
Шевченко Олександр Олександрович

Науковий керівник – к.т.н., доцент
Ліщук Роман Ігорович

Умань – 2024

Зміст

ВСТУП 4

РОЗДІЛ 1 АНАЛІЗ СУЧАСНОГО РОЗВИТКУ ТЕХНОЛОГІЙ РОЗПІЗНАВАННЯ СИМВОЛІВ НА ТЛІ ЗАВАД 7

1.1. Аналіз існуючих методів розпізнавання символів на тлі завад 7

1.2. Аналіз програмних засобів для розпізнавання символів на тлі завад 10

1.3. Постановка задачі 12

РОЗДІЛ 2 АРХІТЕКТУРА ТА МОДЕЛЬ РОЗПІЗНАВАННЯ СИМВОЛІВ НА ТЛІ ЗАВАД 15

2.1. Розробка методу для розпізнавання символів на тлі завад 15

2.2. Обґрунтування вибору типу нейромережі для розпізнавання символів на тлі завад 19

2.3. Модифікація методу подання зображення до нейромережі 33

2.4 Архітектура та модель розпізнавання символів на тлі завад 35

2.5 Складові додатку розпізнавання символів на тлі завад 40

РОЗДІЛ 3 РЕАЛІЗАЦІЯ ЗАСТОСУНКУ ДЛЯ РОЗПІЗНАВАННЯ СИМВОЛІВ НА ТЛІ ЗАВАД НА ОСНОВІ НЕЙРОМЕРЕЖІ 42

3.1. Розробка алгоритму безпосередньо розпізнавання символів на тлі завад 42

3.2. Програні елементи розпізнавання символів на тлі завад 44

3.3. Тестування та аналіз результатів роботи додатку розпізнавання символів на тлі завад 50

ВИСНОВКИ 54

СПИСОК ЛІТЕРАТУРИ 55

ДОДАТКИ 59

ВСТУП

Актуальність теми дослідження. За останні роки комп’ютерні технології змінюються та вдосконалюються з неймовірною швидкістю. Особливо швидко розвиваються портативні пристрої, які наздогнали своїх настільних попередників. Така швидкість розробки обладнання не може не впливати на попит ринку програмного забезпечення. Призначення будь-якого персонального портативного пристрою – автоматизувати певний тип людської діяльності, будь то планування зустрічі чи розрахунок банківської операції.

Окремою сферою діяльності для портативних пристроїв, яка поки що відстає від людини, є природне завдання категоризації та інтелектуальної діяльності. Великих успіхів досягнуто в нових напрямках програмного забезпечення. Голосове керування пристроями вже давно стало стандартною частиною основного програмного забезпечення.

Одним із таких видів діяльності пристрою є завдання ідентифікації. Зі зменшенням розміру та через ергономічні потреби мобільні та портативні пристрої майже втратили свої звичні клавіатури. Однак в ритмі сучасного життя можливість швидкого введення даних в комп'ютер залишається важливою перевагою. Тому потрібні спеціальні програми, здатні швидко реагувати на дії користувача та оцифровувати зрозумілу візуальну інформацію для використання на комп’ютерах.

Процес розпізнавання тексту добре вивчений, з багатьма застосуваннями та робочими прикладами. Проте дослідникам і розробникам все ще ведеться боротьба за пошук нових можливостей щодо якості та швидкості.

Існує досить багато систем загального призначення, які можуть якісно розпізнавати текст і працювати досить швидко, особливо якщо на шляху розпізнавання є проблеми з якістю зображення або розпізнаного об'єкта.

Існує потреба в продукті, який міг би централізовано обробляти вхідну графічну інформацію, розпізнавати сканований текст і обробляти запити користувачів, а з іншого – працювати швидко й ефективно.

Нові розробки мають відповідати двом критеріям: бути повністю адаптивними, а отже універсальними, а також швидко та якісно діяти (чого бракує адаптивним підходам).

Таким чином, застосування сучасних методів штучного інтелекту для вирішення завдань забезпечить подальше застосування інформаційних технологій у сфері розпізнавання символів на тлі перешкод, а також є актуальною темою в дослідженнях нових методів штучного інтелекту.

Мета та мета дослідження. Метою кваліфікаційної роботи є підвищення надійності розпізнавання символів у контексті перешкод для використання штучних нейронних мереж.

Для досягнення поставлених цілей необхідно вирішити такі завдання:

1)Проаналізувати предметну область розпізнавання символів контексті інтерференції;

2)Розглянути існуючі методи вирішення задачі розпізнавання символів розрізі перешкод та обґрунтувати вибір методів, Задовольняти мету даної магістерської кваліфікаційної роботи;

3)Структура нейронної мережі, яка використовується при розробці;

4)Розробка інформаційної технології розпізнавання символів в умовах перешкод;

5)Розробити алгоритм функцій, реалізованих програмним забезпеченням розпізнавання символів у контексті перешкод на основі нейронних мереж;

6)Програмна реалізація інформаційної технології розпізнавання символів на фоні перешкод на основі нейронної мережі;

7)Тестувати та аналізувати програмну реалізацію розпізнавання символів на фоні перешкод на основі нейронних мереж;

Об'єктом дослідження є процес розпізнавання персонажа на фоні перешкод.

Тема дослідження – надійність методів розпізнавання символів і програмних засобів та їх роботи на фоні перешкод на основі нейромережевої структури.

Методи дослідження. У роботі використовуються такі наукові методи дослідження: системний аналіз для аналізу структури інформаційних систем, теорія нейронних мереж для реалізації інформаційних технологій розпізнавання символів у контексті перешкод, методи математичної статистики для розробки процесу з використанням програмних засобів пошуку та розрахувати результати експерименту, орієнтований Програмна реалізація об’єкта програмується.

Практичні наслідки отриманих результатів такі:

1.Розроблено алгоритм розпізнавання символів на основі нейронної мережі на фоні перешкод.

2.Для нейромережевих функцій розроблено алгоритм для підвищення надійності розпізнавання символів в контексті перешкод.

3.Розроблено програмний засіб розпізнавання символів на основі штучної нейронної мережі на фоні перешкод.

РОЗДІЛ 1 АНАЛІЗ СУЧАСНОГО РОЗВИТКУ ТЕХНОЛОГІЙ РОЗПІЗНАВАННЯ СИМВОЛІВ НА ТЛІ ЗАВАД

1.1. Аналіз існуючих методів розпізнавання символів на тлі завад

Найкращі методи розпізнавання зображень/символів використовують оптичні властивості [2, 3]. Тобто символ має декілька властивостей, і система перевіряє наявність цих властивостей у зображення. Цей пошук виконується дуже швидко і має високу метрику точності розпізнавання, але він має кілька явних недоліків [4, 5]. По-перше, для кожного символу та кожного шрифту потрібно знайти та формалізувати десятки або навіть сотні символів, перевіряючи їх на перетини. По-друге, виявлення перешкод або ідентифікація перешкод здійснюється лише за допомогою попередньої обробки, тому зникнення символів призводить до втрати цієї інформації. Загалом, цей спосіб не підходить для розпізнавання рукописних символів і може бути модифікований під поставлене завдання, але показник якості розпізнавання буде дуже низьким.

Для задач розпізнавання образів існує принципово інший підхід.

Для розпізнавання можна використовувати клітинні автомати [5]. Цей автомат є набором локальних правил, які діють на дискретний світ. Ці правила працюють однаково будь-де в дискретному світі та використовують локальне середовище для зміни свого стану. Час в автоматі клітки змінюється в окремих частинах, змінюючи зовнішній вигляд поля клітки. Поле – це N-вимірний простір, розділений на дискретні одиниці за деякими правилами.

Існує багато типів автоматів із кліткою: стандартні автомати з кліткою – засновані на класичному визначенні, наведеному вище; мобільні клітинні автомати – додатково містять активну одиницю, яка змінює свій стан; машина Тюрінга – активна одиниця має кілька станів; альтернативна система – коли клітина зміни блоку тощо.

Ці методи засновані на використанні бінарних зображень і обробці клітинних автоматів для виділення особливостей конкретних символів.

Цей метод вимагає більш ретельної підготовки, вимагає розробки набору систем правил клітинних автоматів і вимагає обробки всіх автоматів одночасно, що вимагає великих обчислень.

Адаптивні методи розпізнавання, засновані на методах інтелектуальної класифікації, є більш загальними.

Одним із найпоширеніших методів є використання нейронних мереж [6].

Нейронні мережі [7-9] – це системи однотипних елементів обробки, що формують вихідні сигнали з вхідних даних. Той факт, що характер мережі змінюється залежно від параметрів елементів, використовується для використання системи класифікації та завдань машинного навчання [10].

Існують методи побудови, методи визначення оптимальних параметрів мережі, таких як кількість шарів, розмір шару, топологія, основні методи представлення вхідного сигналу та їх переваги та недоліки. Одним із найкращих (на практиці) способів представити вхідний сигнал є вибір знімка векторного поля зображення. На відміну від точкових знімків, для цього методу потрібна менша мережа і, отже, менший обсяг Обчислювальний Однак цей метод вимагає більш складної попередньої обробки. Нейронна мережа виконує завдання класифікації самостійно. У багатовимірному просторі нейронна мережа вибирає області, які відповідають деяким екземплярам класу [10, 11].

Головною перевагою нейромережевих структур є їх здатність до навчання. Крім того, вони можуть дуже успішно звертатися до класифікацій об’єктів, які не очікуються під час навчання, оскільки вони наближають об’єкти відповідно до відомих критеріїв.

Недоліками нейромережевих структур є низька точність розпізнавання та високі обчислювальні вимоги (у випадку мережевого моделювання на основі архітектури фон Неймана, тому що зазвичай потреба в швидкості процесора низька, але кількість процесорів важлива). Низька точність також є великим недоліком.

У статті представлено метод побудови та використання нейронних мереж, які ідентифікують не лише первинний вихід нейронної мережі, але й вторинні виходи, які можуть вказувати неправильний напрямок. Такі помилки можна обробляти на етапі постобробки, що зменшить ризик помилок.

Незважаючи на свої недоліки, нейронні мережі є одним із найпотужніших об’єктів для розпізнавання рукописних символів [6]. Структури нейронної мережі дуже лояльно ставляться до незнайомих зображень і намагаються класифікувати їх як найближче знайоме представлення. На практиці людина вводить символ, який дуже схожий на оригінал, але додає деякі помилки нелінійно. У разі розпізнавання друкованого тексту помилка може бути апроксимована як лінійна або нелінійна, якщо не враховувати оптичні перешкоди та шуми. В обох випадках нейронна мережа змогла знайти деякі схожі наближення, тоді як інші методи могли взагалі не розпізнати символи.

Враховуючи наведені методи та їх недоліки, можна зробити висновок, що проблема ідентифікації все ж потребує більш швидкого та точного вирішення. Найпопулярніші та практичні програми використовують регулярність у символах для ідентифікації, що дає високі показники ідентифікації, такі як швидкість або точність, але може стати безсилою у разі перешкод або втрати частини інформації.

Необхідно розробити програму, яка є повністю адаптивною та готовою працювати з будь-яким набором символів в автоматичному режимі. На практиці компанії отримують індивідуальні продукти, в яких програми оптимізовані та доопрацьовані під конкретні завдання. Однак, якщо озирнутися в історію, кожна технологічна революція містить нові можливості для автоматизації. Тому бажаний продукт повинен мати можливість адаптуватися до потреб замовника, будучи при цьому повністю загальним програмним забезпеченням (тобто без втручання програміста).

1.2. Аналіз програмних засобів для розпізнавання символів на тлі завад

Першим аналогом є програма [12]. Це програма для оптичного розпізнавання тексту (OCR - Optical Character Recognition). Він призначений для перетворення сканованих документів, PDF-документів і файлів зображень документів (включно з цифровими фотографіями) у формат. Недоліком такої схеми є низька надійність розпізнавання та висока вартість.

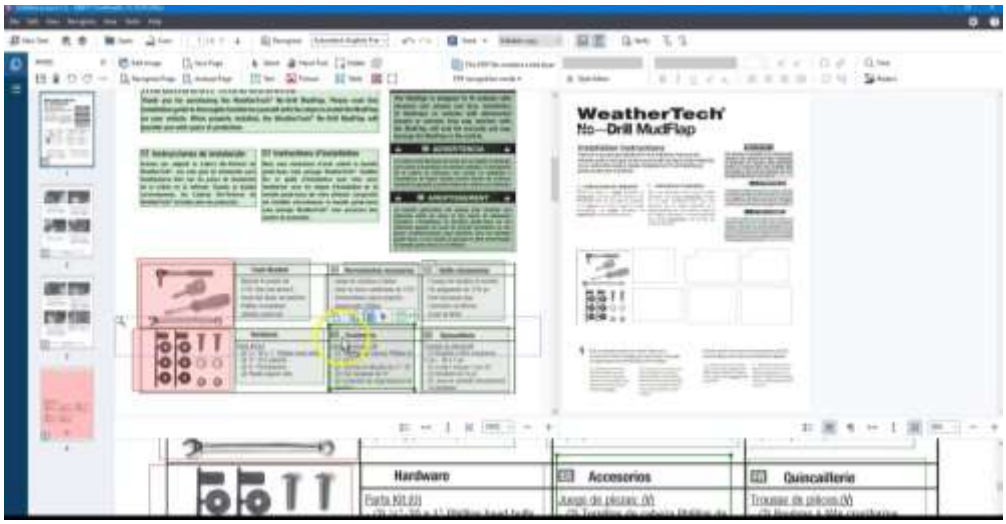


Рисунок 1.1 –Програма ABBYY FineReader

Другим аналогом є програма CueiForm (див. рис. 1.2) [13]. Ця програма була створена компанією Cognitive Technologies для швидкого та високоякісного сканування та розпізнавання документів (роздруківок, фотокопій) як тексту. Щоб розпізнати текст найвищої якості на роздруківках і відсканованому папері, програма використовує спеціальний алгоритм OCR (оптичне розпізнавання символів), який дозволяє легко та швидко визначати мови, шрифти тощо. Недоліком цієї програми є низька надійність розпізнавання, вона може розпізнавати лише друкований текст, а не рукописний.

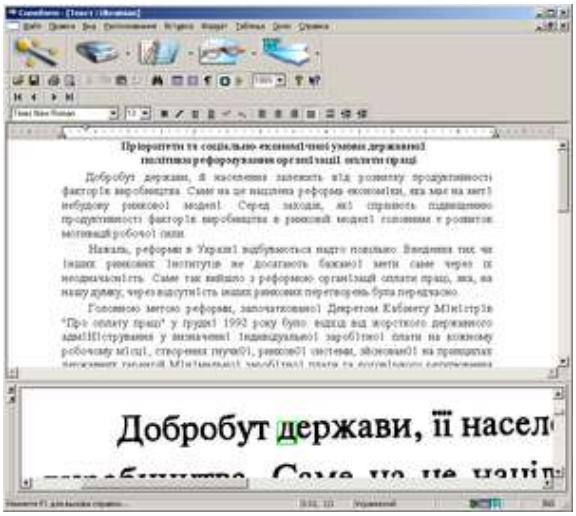


Рисунок 1.2 – Програма CuneiForm

Третьою програмою моделювання є програма MeOCR 1.0 (див. рис. 1.3) [14]. Програма використовує OCR для перетворення сканованих документів у текстові документи, які можна редагувати, і експорту їх у Microsoft Word одним клацанням миші. Його можна використовувати для економії часу, коли вам потрібно перекласти скановані документи в MS Word.

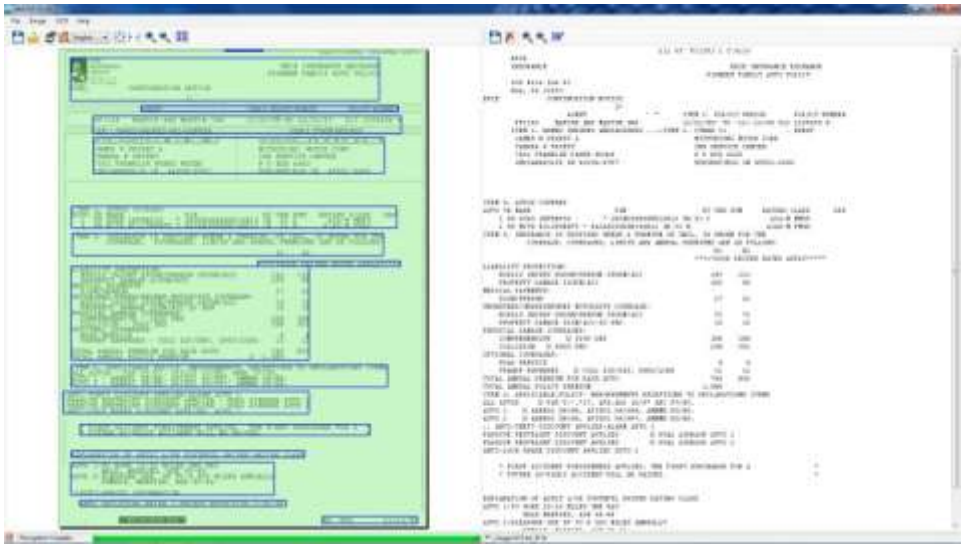


Рисунок 1.3 – Програма MeOCR 1.0

Програма дуже точна і економить час за рахунок зменшення кількості виправлень і необхідних правок. Програма також зберігає формат (більшість програм OCR цього не робить). Він підтримує кілька мов: болгарську, хорватську, чеську, датську, голландську, англійську, естонську, французьку,

німецьку, угорську, італійську, латвійську, литовську, польську, португальську, румунську, російську, сербську, словенську, іспанську, шведську, турецьку, українську. Відскануйте зображення або завантажте зображення з файлу. Підтримуються всі основні формати файлів. BMP, TIFF, JPEG, GIF, PNG. Крім того, програма підтримує колір і може розпізнавати текст на більшості зображень, навіть кольорових. Дозволяє визначити лише частину зображення, вибравши область за допомогою миші. Підтримка пристроїв Twain: зберігайте зображення з будь-якого пристрою Twain, включаючи камери. Також доступна підтримка MS Word: експорт документів MS Word для легкого редагування та форматування. І дозволяє зберігати зображення у форматах BMP, TIFF, JPEG, GIF, PNG. Вбудований текстовий редактор: дозволяє редагувати документи в MeOCR.

Тому в якості розробленої процедури моделювання було обрано процедуру MeOCR [14]. Крім певних переваг, моделювання має і недоліки.

В основному надійність розпізнавання рукописних символів невисока, особливо коли на зображенні є перешкоди, програма на диску велика, програма перевантажена "зайвими функціями", а інтерфейс незручний.

1.3. Постановка задачі

Комп’ютери та комп’ютери розроблені для заміни людей, наприклад, у сферах, де люди виконують свої повсякденні завдання. Сьогодні більшість проблем вирішено, і повсякденні обчислення майже завжди виконуються на ПК. Комп’ютери можуть виконувати прості арифметичні операції швидко та без помилок. Але деякі процеси так і не вирішені до кінця. Особливо це стосується неоднозначних проблем, які вимагають розумних рішень.

Одним із таких завдань є розпізнавання тексту [4]. Розроблено сотні програм, здатних розпізнавати текст. Найкраща розробка використовує дуже прості методи, але завжди вузькоспеціалізована. Тому існують програми, які розробляють систему пошуку областей шаблону для кожного шрифту. Такий

метод працює і діє дуже швидко і ефективно. Але за якістю розпізнавання звичайна людина легко перевершить такі системи.

Людина, використовуючи свій мозок і асоціативну пам'ять, бачить зображення в цілому, використовуючи знання про контекст. Ця додаткова (хоча й незначна) вартість дозволяє досить легко впоратися з розпізнаванням слів і речень, навіть якщо частина зображення пошкоджена. Тому комп'ютеру достатньо знищити 2-3 символи, і результат розпізнавання не задовольнить користувача. Хоча людині легко зрозуміти сенс речення: "Тут якісь дивні символи". Розпізнавання зображень або тексту є або може використовуватися для оптимізації роботи в багатьох сферах. Наприклад, для оцифрування бібліотечних документів, для сканування документів, для автоматичної перевірки. Хоча можна попередньо обробити та підготувати інформацію для сканування, є багато випадків, коли це важко або неможливо використовувати попередню обробку.

Наприклад, автоматизовані системи використовуються для автоматизації реєстрації пасажирів, сканування паспортів для пошуку квитків. Така система має зручний інтерфейс, який не ускладнює процедури для пасажирів, але дозволяє компаніям використовувати менше персоналу. Такі системи успішно працюють у багатьох авіакомпаніях.

Очевидно, що для ефективності дуже важливо, щоб якість системи була якомога вищою, оскільки інформацію неможливо ідентифікувати під час її читання, інакше можливі наслідки помилки можуть призвести до того, що клієнт отримає не той квиток. В інших випадках такі випуски можуть стосуватися цінних паперів і мати більш серйозні наслідки.

Однак ми зосередимося на загальних системах розпізнавання зображень, а також тексту, які виконуватимуть розпізнавання змістовного тексту, але також будуть готові розпізнавати окремі семантично нерелевантні символи.

Така система є більш загальною і може бути налаштована під конкретного замовника, крім того, вона має можливість навчитися розпізнавати друкований текст і Розпізнавання рукописного тексту.

Магістерська робота в цьому розділі стосується недостатньої надійності сучасних процесів розпізнавання символів у контексті перешкод. Розглянуто основні методи та програмні засоби розпізнавання символів у контексті перешкод та визначено наступні недоліки:

1) Надійність розпізнавання недостатня, тобто порівняно небагато символів правильно розпізнається за наявності перешкод у зображенні символу, 2) Низька швидкість розпізнавання, що пов’язано з послідовним характером порівняння символів із посилання.

Для усунення цих недоліків запропоновано метод розпізнавання символів на фоні перешкод на основі штучної нейронної мережі. На основі цього методу необхідно розробити інформаційну технологію, впровадити ІТ-програмне забезпечення, перевірити виконання вимог до програмних засобів та усунути встановлені недоліки.

РОЗДІЛ 2 АРХІТЕКТУРА ТА МОДЕЛЬ РОЗПІЗНАВАННЯ СИМВОЛІВ НА ТЛІ ЗАВАД

2.1. Розробка методу для розпізнавання символів на тлі завад

Розпізнавання символів є дуже складним завданням як в теоретичному, так і в практичному сенсі, хоча багато організмів, включаючи людину, можуть впоратися з нею швидко і досить точно. Для ефективного виконання цього процесу досить складно створити штучну систему і реалізувати її технічно. Розпізнавання образів – це порівняння властивостей об’єкта із зображенням об’єкта.

Прикладом розпізнавання зображень може бути система розпізнавання тексту та окремих символів, біометричних параметрів людей, штрих-кодів, номерів транспортних засобів тощо. Система розпізнавання тексту передбачає, що на вході є зображення з текстом (в заданому графічному форматі). Текст повинен генеруватися на виході системи. Загальна архітектура такої системи складається з наступних послідовних операцій [3, 5]:

1. Коли вводиться зображення, спочатку потрібно видалити шум і зменшити його до форми, яка дозволяє ефективний вибір символів та їх ідентифікацію.

2. Система повинна сегментувати зображення на текстові блоки відповідно до характеристик зображення.

3. Зазвичай система повинна розкладати блоки тексту на слова, а слова – на символи, оскільки розпізнавання символів простіше, ніж розпізнавання цілих слів.

4. Після цього система виконує певний алгоритм ідентифікації. Саме розпізнавання символів може виконуватися різними специфічними методами. Ви можете порівнювати з шаблонами, шукати фрагменти інформації, очікувані для певного символу, використовувати нейронні мережі тощо. [15, 16].

Після ідентифікації вони отримують набір символів, з яких формується текст.

Існує кілька основних систем розпізнавання тексту [4, 16]. Принципи роботи таких систем базуються на різноманітних стратегіях, але в цілому алгоритми в загальному вигляді базуються на формулюванні гіпотез і їх подальшій перевірці. Крім того, швидкість алгоритму залежить від порядку припущень, тому часто важливо знайти найкращий порядок припущень на основі даних, отриманих із зображень.

OCR зазвичай застосовується до растрових зображень сторінки. Тим часом більшість систем створюють шаблони для різних шрифтів і контурів. Програма розпізнає шрифт і розпізнає його на основі шаблонів. У деяких випадках програмне забезпечення використовує числові значення деяких символів для визначення нових шрифтів [7].

Найпопулярніші програми розпізнавання тексту засновані на пошуку певних характеристик символів. Така система складається із сотень алгоритмів, кожен з яких шукає ідентифікаційний символ для певного символу. Такі системи іноді використовують принцип нечіткої множини для визначення ймовірності розпізнавання символу. Однак такі системи мають ряд недоліків. Відсутність одного пікселя або наявність регулярного чи нерегулярного шуму може значно погіршити якість розпізнавання. Нейронна мережа може бути класифікатором у системі розпізнавання тексту. Цей класифікатор можна навчити, регулюючи коефіцієнти на елементах мережі, щоб, залежно від вхідних даних, бажаний результат з’являвся на виході. Нейронні мережі можуть успішно використовуватися в системах розпізнавання образів. Суть мережі полягає в тому, що вона виділяє регіони в багатовимірному просторі всіх можливих вхідних даних. Наприклад, кожному регіону може відповідати символ. Якщо після навчання точка з цього простору прикладається до входу мережі (тобто будь-якого вхідного сигналу), мережа намагатиметься визначити, до якого регіону належить сигнал. В результаті мережа здатна реагувати навіть на вхідні сигнали, з якими мережа ще не стикалася, а результатом роботи буде область, найближча до заданої точки. Тому точність мережі фактично залежить від якості навчання [10].

Одним з недоліків нейронних мереж є їх розмір. Зазвичай потрібна досить велика структура мережі. Однак ця складність зумовлена традиційними методами обробки інформації. Сьогодні більшість комп'ютерів складається з кількох дуже потужних процесорів, які можуть дуже швидко виконувати атомарні арифметичні операції. Нейронні мережі, по суті, складаються з великої кількості простих елементів, які просто виконують дію суматора з пороговою функцією. Однак таких елементів багато. Якщо порівняти біологічну нейронну мережу людини з сучасним процесором, людина виконує прості математичні операції, такі як множення або ділення, дуже повільно, хоча процесор виконує їх за мільярдні частки секунди. З іншого боку, достатньо 0,2 секунди, щоб зрозуміти, що картина написана в стилі кубізму, що для бінарного процесора є майже непосильним завданням.

Більш простим варіантом реалізації нейронної мережі є моделювання її на звичайному процесорі. Однак процесори обробляють інформацію послідовно, тому, хоча кожен елемент мережі незалежний і всі обчислення можуть виконуватися паралельно, час, необхідний для обробки одного вхідного сигналу, пропорційний розміру мережі [9, 11].

Існує багато типів нейронних мереж, які поділяються на одношарові та багатошарові.

Хоча один нейрон може виконувати найпростіші процедури розпізнавання, потужність нейронних обчислень походить від з'єднань нейронів у мережі. Найпростіша мережа складається з набору нейронів, які утворюють шар, як показано на правій стороні рисунка 2.1. Вершини кола зліва використовуються лише для призначення вхідного сигналу. Вони не виконують обчислень і тому не вважаються шаром. З цієї причини вони позначені маленькими кружками, щоб відрізнити їх від обчислювальних нейронів, позначених великими кружками. Кожен елемент у вхідному наборі X пов'язаний з кожним штучним нейроном за допомогою окремої ваги, і кожен нейрон виводить зважену суму своїх входів у мережу.

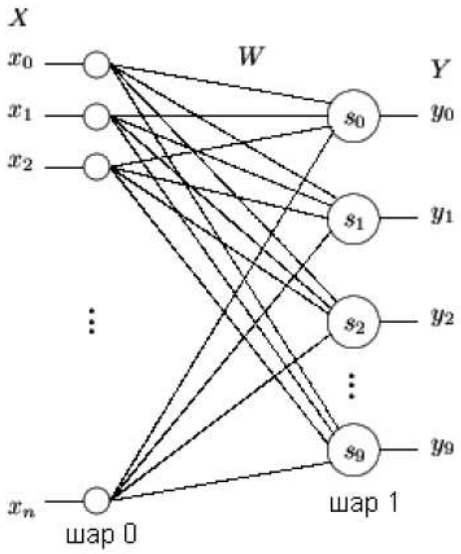


Рисунок 2.1 – Одношарова нейромережа

Великі та складні нейронні мережі зазвичай мають велику обчислювальну потужність. Хоча були створені мережі будь-якої мислимої конфігурації, ієрархічна організація нейронів імітує ієрархічну структуру певних областей мозку. Такі багаторівневі мережі мають більші можливості, ніж одношарові мережі.

Багатошарові мережі можуть бути сформовані з каскадних шарів. Вихід одного шару є входом наступного шару. Аналогічна мережа показана на малюнку. 2.2 Лише коли функція активації між рівнями є лінійною, багаторівнева мережа може підвищити обчислювальну потужність порівняно з одноріневою мережею. Обчислення результату шару складається з множення вхідного вектора на першу вагову матрицю, а потім множення результуючого вектора на другу вагову матрицю (якщо немає нелінійної функції активації).

Оскільки множення матриці є асоціативним, то це показує, що двошарова лінійна мережа еквівалентна одному шару з ваговою матрицею, яка дорівнює добутку двох вагових матриць. Тому будь-яку багатошарову лінійну мережу можна замінити еквівалентною одноріневою мережею, але обчислювальна потужність одношарової мережі досить обмежена. Тому для розширення можливостей мережі в порівнянні з одноріневими мережами потрібні нелінійні функції активації.

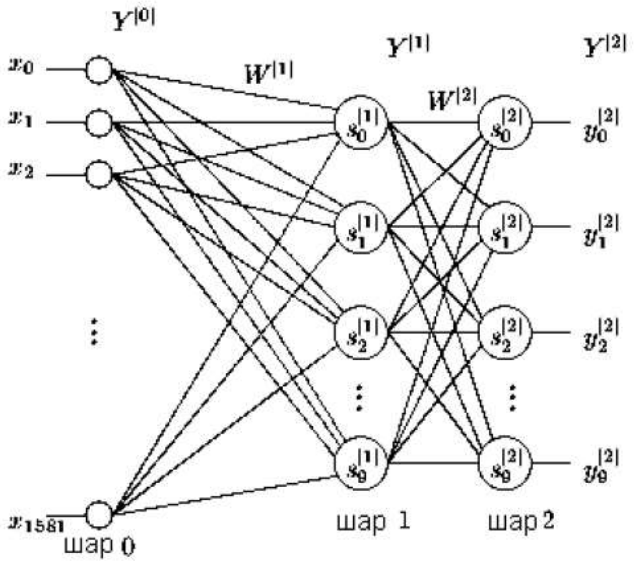


Рисунок 2.2 – Багатошарова нейромережа

2.2. Обґрунтування вибору типу нейромережі для розпізнавання символів на тлі завад

Нейронні мережі справді є найкращими кандидатами для реалізації адаптивних рішень проблем класифікації. Однак для ефективного вирішення поставленої задачі необхідно вибрати певний тип мережі та проаналізувати існуючі варіанти реалізації нейронних мереж.

Для огляду було обрано чотири популярні варіанти реалізації нейронних мереж: нейронні мережі із зворотним поширенням помилок, нейронні мережі Хебба, нейронні мережі Хопфілда та нейронні мережі Хеммінга.

Аналіз кількох варіантів дозволить більш детально розглянути математичну основу процесу навчання нейронної мережі та прийняти правильні рішення щодо використання тієї чи іншої реалізації.

Нейромережі із зворотним поширенням помилки.

Коли в мережі є тільки один рівень, його та навчені викладачами алгоритми очевидні, оскільки відомий правильний початковий стан нейронів в одному шарі, а налаштування синаптичних зв’язків здійснюється в напрямку мінімізації помилок у виході. мережі. Наприклад, за таким принципом побудований алгоритм навчання одношарового персептрона.

У багатошаровій мережі оптимальне вихідне значення нейронів у всіх шарах, крім останнього, зазвичай невідоме, і більше неможливо навчити персептрон із двома або більше шарами, тільки вихід штучної нейронної мережі. У розглянутій вище мережі відсутні зворотні зв'язки, тобто зв'язки від виходу певного рівня до входу того ж рівня або попереднього рівня. Такі спеціальні мережі називаються мережами без зворотного зв'язку або мережами прямого поширення. У мережах без зворотного зв'язку пам'ять відсутня, а їх вихід повністю визначається поточним входом і значеннями ваги.

Одним із варіантів вирішення цієї проблеми є розробка набору вихідних сигналів, відповідних вхідним сигналам кожного рівня ШНМ, звичайно, це дуже трудомістка операція і не завжди здійсненна [10, 11]. Другий варіант - це динамічне коригування синаптичних ваг, під час якого зазвичай вибирається найслабший зв'язок і вносяться невеликі зміни в один або інший бік, і лише ті зміни, які призводять до зменшення помилки виведення, зберігаються по всій мережі. Третій варіант полягає в поширенні сигналу помилки від виходу ШНМ до його входу в напрямку, протилежному прямому поширенню сигналу в нормальному режимі роботи. Цей алгоритм навчання ШНМ називається процесом зворотного поширення.

Відповідно до методу найменших квадратів цільовою функцією мінімізації похибки ШНМ є:

$$E(w) = \frac{1}{2}\sum_{j,p}(y_{j,p}^{(N)} - d_{j,p})^2 \qquad (2.1)$$

додати всі нейрони та всі зображення до вихідного рівня, обробляється мережею. Він мінімізується методом градієнтного спуску, тобто ваговий коефіцієнт регулюється наступним чином:

$$\Delta w_{ij}^{(n)} = -\mu\frac{\partial E}{\partial w_{ij}} \qquad (2.2)$$

Тут коефіцієнт швидкості навчання:

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial y_j}\frac{dy_j}{ds_j}\frac{\partial s_j}{\partial w_{ij}} \tag{2.3}$$

Тут - вихід нейрона - зважена сума його вихідних сигналів – параметрів функції активації. Оскільки множник є похідною функції по її аргументам, то похідну функції активації необхідно визначати по всій осі абсцис. У зв'язку з цим унімодальні функції та інші неоднорідні функції активації не підходять для розглянутої штучної нейронної мережі. Вони використовують функції згладжування, такі як гіперболічний тангенс або класична сигмоподібна з експонентою. У випадку гіперболічного тангенса:

$$\frac{dy_j}{ds_j} = 1 - s^2 \tag{2.4}$$

Третій множник, очевидно, дорівнює нейрону попереднього шару. Перший множник легко розкладається наступним чином:

$$\frac{\partial E}{\partial y_i} = \sum_k \frac{\partial E}{\partial y_k}\frac{\partial y_k}{\partial s_k}\frac{\partial s_k}{\partial y_i} = \sum_k \frac{\partial E}{\partial y_k}\frac{\partial y_k}{\partial s_k} w_{jk}^{(n+1)} \tag{2.5}$$

Сумування по виконується серед нейронів шару. Ввівши нову змінну

$$\delta_j^{(n)} = \frac{\partial E}{\partial y_j}\frac{\partial y_j}{\partial s_j} \tag{2.6}$$

отримаємо рекурсивну формулу

$$\delta_j^{(n)} = \left[\sum_k \delta_j^{(n+1)} w_{jk}^{(n+1)}\right]\frac{\partial y_j}{\partial s_j} \tag{2.7}$$

Для вихідного шару

$$\delta_l^{(N)} = \left(y_l^{(N)} - d_l\right)\frac{dy_l}{ds_l} \tag{2.8}$$

Тепер можемо записати (2.6) в розкритому вигляді:

$$\Delta w_{ij}^{(n)} = -\mu\delta_j^{(n)} y_i^{(n-1)} \tag{2.9}$$

Іноді, щоб надати процесу корекції ваги деяку інерцію для згладжування різких стрибків при русі по поверхні цільової функції, доповнюють (2.9) значенням зміни ваги на попередній ітерації

$$\Delta w_{ij}^{(n)}(t) = -\mu(\rho \Delta w_{ij}^{(n)}(t-1) + (1-\rho)\delta_j^{(n)} y_i^{(n-1)}) \qquad (2.10)$$

Таким чином, повний алгоритм навчання ШНМ з використанням процесу зворотного поширення помилок виглядає наступним чином:

1. Нанесіть одне з можливих зображень у нормальному режимі роботи нейронної мережі на вхід мережі та обчисліть значення останнього при поширенні сигналу від входу до виходу

$$s_j^{(n)} = \sum_{i-0}^{M} y_i^{(n-1)} w_{ij}^{(n)} \qquad (2.11)$$

$$y_j^{(n)} = f\left(s_j^{(n)}\right), \text{де } f(x) - \text{ сигмоід} \qquad (2.12)$$

$$y_q^{(0)} = I_q \qquad (2.13)$$

2. Розрахувати для вихідного шару по формулі (2.8). Розрахувати по формулі (2.9) або (2.10) зміни ваг шару .

3. Розрахувати по формулах (2.6) і (2.9) (або (2.8) і (2.10)) відповідно.

4. Скоригувати всі ваги в ШНМ

$$w_{ij}^{(n)}(t) = w_{ij}^{(n)}(t-1) + \Delta w_{ij}^{(n)}(t) \qquad (2.14)$$

5. Якщо помилка мережі серйозна, перейдіть до кроку 1. В іншому випадку кінець.

На кроці 1 мережа чергує всі навчальні зображення у випадковому порядку, щоб мережа не залишила деякі зображення, як запам’ятовує інші.

З рівняння (2.12) видно, що коли початкове значення наближається до нуля, ефект навчання значно зменшується. Використовуючи бінарні вхідні вектори, в середньому половина вагових коефіцієнтів не буде скоригована, тому бажано перемістити діапазон можливих значень виходу нейрона [0,1] до межі [-0,5,+0,5], тобто. шляхом застосування логістичної функції до простої

модифікації для досягнення. Наприклад, сигма з експонентою перетворюється у форму

$$f(x) = -0.5 + \frac{1}{1 + e^{-ax}} \tag{2.15}$$

У процесі навчання може виникнути ситуація, коли велике позитивне або негативне значення вагового коефіцієнта перемістить робочу точку на сигмовиді багатьох нейронів в область насичення. Відповідно до (2.7) і (2.8), мале значення похідної логістичної функції призведе до зупинки навчання, таким чином паралізуючи ШНМ. По-друге, застосування градієнтного спуску не гарантує, що буде знайдено глобальний мінімум цільової функції, а не локальний мінімум. Ця проблема пов'язана з іншою проблемою - вибором значення швидкості навчання. Доказ конвергенції навчання під час зворотного поширення базується на похідних, тобто збільшенні ваг, тому швидкість навчання має бути нескінченно малою, але в цьому випадку швидкість навчання буде неприйнятно низькою. З іншого боку, занадто велика корекція ваги призведе до постійної нестабільності процесу навчання. Тому в якості η зазвичай вибирають число менше 1, але не надто мале, наприклад 0,1, яке можна поступово зменшувати в процесі навчання. Крім того, щоб виключити випадкові попадання в локальні мінімуми, іноді, після стабілізації значення вагових коефіцієнтів, η короткочасно сильно збільшується, щоб почати градієнтний спуск з нової точки. Якщо багаторазове повторення цього процесу переводить алгоритм у той самий стан ШНМ, ми можемо з більшою чи меншою впевненістю сказати, що знайдено глобальний максимум, і нічого більше.

Існує інший спосіб усунення локальних мінімумів, який також усуває параліч ШНМ, і він передбачає застосування випадкових ШНМ.

Алгоритм навчання нейронної мережі з використанням процесу зворотного поширення означає, що існують деякі зовнішні зв’язки, які надають мережі вхідні та цільові вихідні зображення. Алгоритми, які використовують схожі концепції, називаються алгоритмами навчання з наставником. Для його

успішної роботи необхідно, щоб експерти на попередньому етапі створили еталонні виходи для кожного вхідного зображення. Наприклад, навчання людського мозку, на перший погляд, відбувається без вчителя: зорові, слухові, тактильні та інші рецептори сприймають інформацію ззовні і здійснюють своєрідну самоорганізацію всередині нервової системи. Проте беззаперечно, що в житті є чимало вчителів – як у прямому, так і в переносному значенні – які координують зовнішні впливи.

Основна характеристика навчання без вчителя – його "самостійність". Процес навчання з викладачем передбачав регулювання ваги синапсів. Деякі алгоритми змінюють структуру мережі, тобто взаємозв'язок нейронів і навіть їх кількість, це перетворення називається самоорганізацією. Зрозуміло, що коригування синапсу може здійснюватися лише на основі інформації, наявної в нейроні, а саме його стану та вагових коефіцієнтів, які вже існують. На основі цих міркувань і за аналогією з відомим принципом самоорганізації нейронів був створений алгоритм навчання Хебба.

Нейромережі Хебба.

За Хеббом, нервові клітини головного мозку з'єднані між собою великою кількістю прямих і зворотних збудливих зв'язків, утворюючи нейронну мережу. Кожен нейрон просторово-часово підсумовує сигнали від пошкодженого нейрона, щоб визначити потенціал через його мембрану. Нейрони спрацьовують, коли потенціал мембрани перевищує порогове значення. Нейрони вогнестійкі та стійкі до втоми. Ефективність зв'язків може змінюватися під час роботи мережі, збільшуючись між порушеними нейронами. Це змушує нейрони об’єднуватися в групи клітин – ті, які найчастіше порушуються разом – і змушує групи клітин відокремлюватися одна від одної. Коли ансамбль порушується, він повністю збуджений. Різні ансамблі можуть перекриватися: той самий нейрон може входити до різних ансамблів. Електрична активність мозку зумовлена послідовним збудженням єдиного цілого [10, 15].

Метод вивчення сигналу Хебба полягає у зміні вагових коефіцієнтів відповідно до наступних правил:

$$w_{ij}(t) = w_{ij}(t-1) + \alpha y_i^{(n-1)} y_j^{(n)} \qquad (2.16)$$

Існує також і диференціальний метод навчання Хебба:

З рівняння (2.17) можна побачити, що синапси, що з'єднують ці нейрони, вихід яких найбільше змінюється в напрямку збільшення, навчаються найбільше.

$$w_{ij}(t) = w_{ij}(t-1) + \alpha\left[y_i^{(n-1)}(t) - y_i^{(n-1)}(t-1)\right]\left[y_j^{(n)}(t) - y_j^{(n)}(t-1)\right] \qquad (2.17)$$

З рівняння (2.17) можна побачити, що синапси, що з'єднують ці нейрони, вихід яких найбільше змінюється в напрямку збільшення, навчаються найбільше.

Повний алгоритм навчання за наведеною вище формулою виглядатиме так:

1. Під час фази ініціалізації всім вагам присвоюються невеликі випадкові значення.

2. Вхідне зображення подається на вхід мережі, сигнал збудження поширюється по всіх шарах за принципом класичної прямої мережі, тобто для кожного нейрона обчислюється зважена сума його входів і потім застосовується функція активації (передачі) нейрона, результатом є її початкове значення.

3. Змінити ваговий коефіцієнт відповідно до початкового значення нейрона, отриманого за формулою (2.16) або (2.17).

4. Цикл - перехід до кроку 2, поки вихідне значення мережі не стабілізується до заданої точності.

Використання цього нового методу визначення завершення навчання відрізняється від того, що використовується для мереж зворотного поширення, оскільки значення налаштування синапсів практично необмежене. На другому кроці циклу чергуйте всі зображення у вхідному наборі.

Слід зазначити, що через випадковий розподіл ваг на фазі ініціалізації тип відповіді на кожен клас вхідних зображень невідомий заздалегідь і представлятиме будь-яку комбінацію станів нейрона на вихідному рівні. У той же час мережа здатна узагальнювати подібні зображення, відносячи їх до одного класу. Тестування навченої мережі дозволяє визначити топологію класів на вихідному рівні. Щоб відповіді навченої мережі було зручно представлено, мережу можна доповнити шаром, наприклад, за алгоритмом навчання одношарового персептрона, який необхідно змусити відображати початкові відповіді мережі в необхідних зображеннях. .

Інший алгоритм навчання без вчителя - алгоритм Кохонена - передбачає коригування синапсів на основі значень з попередніх ітерацій.

$$w_{ij}(t) = w_{ij}(t-1) + \alpha\left[y_i^{(n-1)} - w_{ij}(t-1)\right] \qquad (2.18)$$

Як видно з наведеної вище формули, навчання спрощено, щоб мінімізувати різницю між вхідним сигналом нейрона від виходу нейронів попереднього шару та ваговими коефіцієнтами його синапсів [8, 9].

Повний алгоритм навчання має приблизно таку саму структуру, що й метод Хебба, але на кроці 3 із усього шару вибирається нейрон, значення синапсу якого максимально подібне до вхідного зображення, а ваги регулюються відповідно до рівняння (2.18) лише для це. Це так зване схвалення може супроводжуватися гальмуванням усіх інших нейронів у шарі та введенням вибраних нейронів у насичений стан. Такі нейрони можна вибрати, наприклад, шляхом обчислення скалярного множення вектора вагових коефіцієнтів і вектора вхідних значень. Нейрон-переможець дає найбільший результат.

Іншим варіантом є обчислення відстані між цими векторами в p-вимірному просторі, де p – розмір вектора.

$$D_j = \sqrt{\sum_{i=0}^{p-1}(y_i^{(n-1)} - w_{ij})^2} \qquad (2.19)$$

де індекс нейрона n-го шару, і – індекс підсумовування нейрона (n-1)-го шару, вага синапсів, що з’єднують нейрони; вихід нейрона (n-1)-го шару – n- й шар вхідне значення. Немає необхідності знаходити корені в рівнянні (2.19), оскільки це просто інша відносна оцінка.

У цьому випадку "перемагає" нейрон з найменшою відстанню. Іноді нейрони примусово виключаються з розгляду через занадто часту сертифікацію, щоб "зрівняти" права всіх нейронів у шарі. Найпростіший варіант цього алгоритму полягає в придушенні тільки переможних нейронів.

При використанні навчання за алгоритмом Кохонена одним з підходів є нормалізація вхідного зображення та - на етапі ініціалізації - нормалізація початкових значень вагових коефіцієнтів.

$$x_i = x_i / \sqrt{\sum_{j=0}^{n-1} x_j^2} \qquad (2.20)$$

де – компоненти вектора вхідного зображення або вектора вагових коефіцієнтів, а n – його розмірність. Це дозволяє скоротити тривалість процесу навчання.

Ініціалізація вагових коефіцієнтів випадковими значеннями може призвести до злиття різних класів, що відповідають щільно розподіленим вхідним зображенням, або, навпаки, розбиття на додаткові підкласи у випадку близьких зображень одного класу. Щоб уникнути цього, використовуються опуклі методи комбінування. Його суть зводиться до того, що вхідне нормалізоване зображення трансформується:

$$x_i = \alpha(t) x_i + \left(1 - \alpha(t)\right) \frac{1}{\sqrt{n}} \qquad (2.21)$$

де коефіцієнти змінюються від нуля до одиниці під час процесу навчання, тому спочатку майже ідентичні зображення застосовуються до входу мережі, а з часом вони все більше зближуються до оригінального зображення.

Коефіцієнт ваги встановлюється на етапі ініціалізації як

$$w_o = \frac{1}{\sqrt{n}} \tag{2.22}$$

де n - розмірність вектора ваги для нейронів ініціалізованого шару.

На основі описаного вище методу будується особливий тип нейронної мережі – так звана самоорганізована структура із синаптичним налаштуванням – як самоорганізована карта ознак. Для них після вибору нейрона j з найменшою відстанню від n-го шару (2.19) за формулою (2.18) навчається не тільки цей нейрон, а й його сусіди по R. Значення R дуже велике на першій ітерації, щоб усі нейрони були навчені, але з часом воно зменшується до нуля. Тому чим ближче до кінця навчання, тим точніше можна визначити групу нейронів, що відповідає кожному класу зображень.

Нейромережі Хопфілда

Серед різноманітних конфігурацій штучних нейронних мереж деякі конфігурації класифікуються за принципом навчання. У такій мережі ваги синапсів обчислюються тільки один раз перед початком роботи мережі на основі інформації оброблених даних, і все навчання мережі зводиться до цього обчислення [11]. З одного боку, представлення попередньої інформації можна розглядати як допомогу вчителю, але з іншого боку, мережа насправді просто запам’ятовує зразки до того, як на її вхід надходять реальні дані, і не може змінити свою поведінку, тому ми можемо говорити про зв'язок зі "світом" (вчителем) зв'язки петлі зворотного зв'язку не доведені. Серед мереж зі схожою логікою роботи найбільш відомими є мережа Хопфілда і мережа Хеммінга, які зазвичай використовуються для організації асоціативної пам'яті.

Структурна схема мережі Хопфілда показана на малюнку 1. 2.3 Він складається з одного шару нейронів, кількість яких є одночасно кількістю входів і виходів мережі. Кожен нейрон з’єднаний з усіма іншими нейронами синапсами, а також має вхідний синапс, через який надходять сигнали. Вихідний сигнал зазвичай генерується на аксоні.

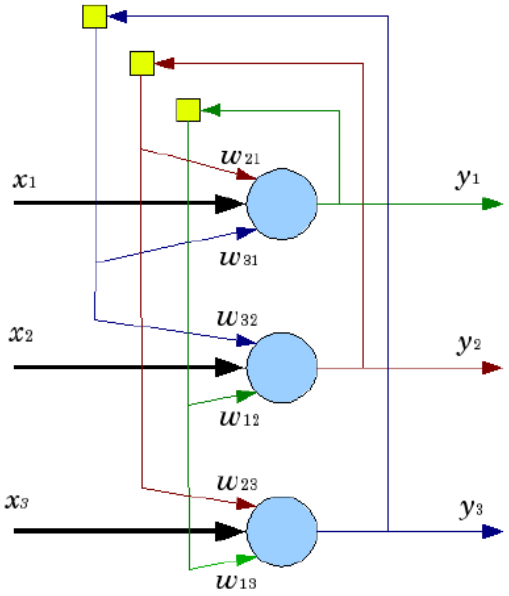


Рисунок 2.3 –Схема нейромережі Хопфілда

Наприклад, якщо сигнал являє собою якесь зображення, то, відобразивши дані з мережевого виходу в графічній формі, ви зможете побачити картинку, яка точно відповідає одному з прикладів (у разі успіху) або " вільна імпровізація" мережі (в невдачі).

Під час фази ініціалізації мережі синаптичні ваги встановлюються наступним чином:

$$w_{ij} = \begin{cases} \sum_{k=0}^{m-1} x_i^k x_j^k, i \neq j \\ 0, \qquad i = j \end{cases} \tag{2.23}$$

Алгоритм роботи мережі такий (p – кількість ітерацій):

1. На вхід мережі подається невідомий сигнал. Насправді його введення здійснюється шляхом безпосередньої установки величини аксона:

$$y_i(0) = x_i, i = 0, \dots, n-1 \tag{2.24}$$

Тому вказівка вхідної синаптичної мережі в явному вигляді є суто умовною. Нуль у правій дужці означає нульову ітерацію в мережевому циклі.

2. Розрахувати новий стан нейрона

$$s_i(p+1) = \sum_{i=0}^{n-1} w_{ij} y_i(p), j = 0, \dots, n-1 \tag{2.25}$$

і нові значення аксонів

$$y_i(p+1) = f\left[s_j(p+1)\right] \qquad (2.26)$$

де активаційна функція у вигляді стрибка, наведена на рис. 2.4.

3. Перевірте, чи змінилося початкове значення аксона на останній ітерації. Якщо так - переходимо до пункту 2, інакше (якщо вихід стабілізувався) - закінчуємо. Тим часом вихідний вектор є зразком, який найкраще відповідає вхідним даним.

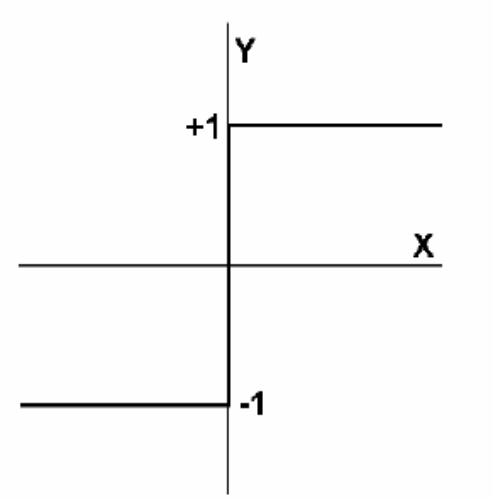


Рисунок 2.4 – Стрибкоподібна функція активації

Іноді мережа не розпізнає і виводить зображення, якого не існує. Це пов’язано з обмеженими можливостями мережі. Для мереж Хопфілда число m образів пам'яті не повинно перевищувати значення, приблизно рівне 0,15 * n. Крім того, якщо два зображення A і B дуже схожі, вони можуть викликати взаємну кореляцію в мережі, тобто поява вектора A на вході мережі спричинить появу вектора B на її виході, і навпаки.

Нейромережі Хеммінга.

Коли мережі не потрібно відправляти вибірку явно, тобто достатньо отримати номер вибірки, і мережа Хеммінга успішно реалізує асоціативну пам'ять. Порівняно з мережею Хопфілда, мережа характеризується меншою вартістю пам'яті та обчислювальними зусиллями, що видно з її структури.

Мережа Хеммінга (рис. 2.5) складається з двох шарів. Перший і другий шари мають по m нейронів, де m – кількість вибірок. Кожен нейрон першого шару має n синапсів, підключених до входу мережі (утворюючи уявний

нульовий шар). Нейрони другого шару з'єднані один з одним за допомогою гальмівних (негативний зворотний зв'язок) синаптичних зв'язків. Кожен нейрон має унікальні синапси позитивного зворотного зв'язку, з'єднані з його аксонами.

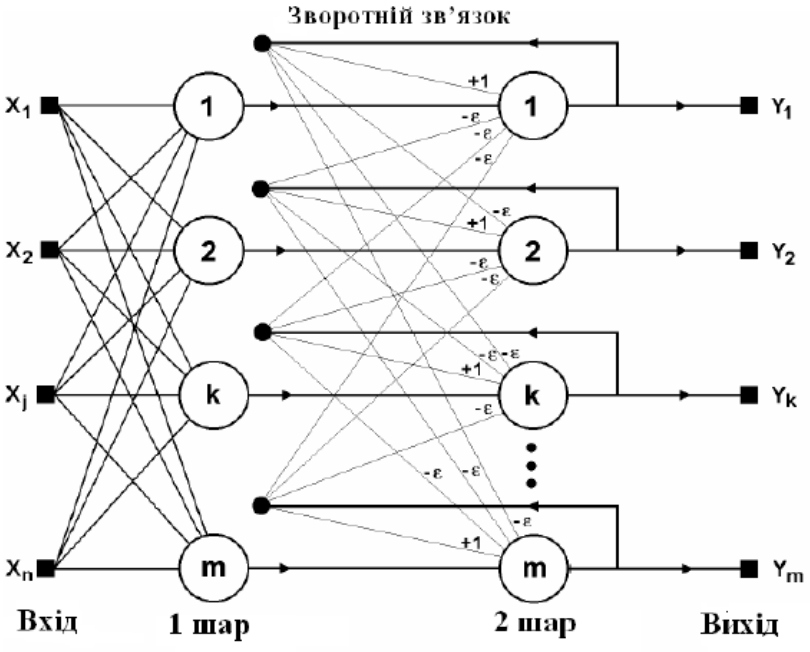


Рисунок 2.5 – Схема нейромережі Хеммінга

Ідея мережі полягає в тому, щоб знайти відстань Хеммінга від тестованого методу до всіх зразків. Відстань Хеммінга – це кількість бітів, які відрізняються у двох двійкових векторах. Мережа повинна вибрати вибірку з найменшою відстанню Хеммінга від невідомого вхідного сигналу, в результаті чого активується лише один вихід мережі, що відповідає цій вибірці.

Під час фази ініціалізації вагам першого рівня та порогу функції активації присвоюються наступні значення:

$$w_{ik} = \frac{x_i^k}{2}, i = 0, \dots, n-1, k = 0, \dots, m-1 \quad (2.27)$$

$$T_k = \frac{n}{2}, k = 0, \dots, m-1 \quad (2.28)$$

Ваговий коефіцієнт другого шару гальмівних синапсів дорівнює певній величині. Синапс нейрона, з'єднаний з його аксоном, має вагу +1.

Алгоритм роботи мережі Хеммінга виглядає наступним чином:

1. Невідомий вектор X подається на вхід мережі, який обчислює стан нейронів першого шару (верхній індекс у дужках вказує кількість шарів):

$$y_j^{(1)} = s_j^{(1)} = \sum_{i=0}^{n-1} w_{ij} x_i + T_j \quad , j = 0, \dots, m-1 \tag{2.29}$$

Після цього отриманим значенням ініціалізуються значення аксонів другого шару:

$$y_j^{(2)} = y_j^{(1)}, j = 0, \dots, m-1 \tag{2.30}$$

і значення їх аксонів:

$$y_j^{(2)}(p+1) = f\left[s_j^{(2)}(p+1)\right], j = 0, \dots, m-1 \tag{2.32}$$

Функція активації має вигляд порогу, а значення F має бути достатньо великим, щоб будь-яке можливе значення параметра не призводило до насичення.

3. Перевірте, чи змінилися вихідні дані нейронів другого шару під час останньої ітерації. Якщо так, перейдіть до кроку 2. В іншому випадку кінець.

З оцінки алгоритму видно, що роль першого рівня досить умовна: значення його вагового коефіцієнта використовується один раз на кроці 1, і мережа більше не посилається на нього, тому перший рівень може бути повністю виключено з мережі (замінено матрицею вагових коефіцієнтів).

Розглянувши чотири типи мереж, ми виберемо ту, яка оптимізована для реалізації в системі розпізнавання образів. Для цього ми розглянемо найважливіші для нас критерії.

Оскільки система має працювати якомога ефективніше, потрібен мінімальний розмір мережі, оскільки навчання мережі є процесом, що вимагає обчислень.

Перший критерій – швидкість. Навчання мережі має вимагати якомога менше операцій. Для підвищення продуктивності можна додати додаткову вимогу: можливість тренуватися на цілочисельній арифметиці. Однак такий підхід знижує точність рішення, що негативно позначається на другому

критерії: якості розпізнавання. Отже, давайте просто зосередимося на ефективності з точки зору найменшої операції ділення, множення та додавання (віднімання, оскільки ці дві операції еквівалентні з точки зору швидкості з точки зору комп’ютера) [11].

За першим критерієм найкращим типом мережі є традиційна мережа зворотного поширення помилок. Це демонструє той факт, що навчання можна оптимізувати, вибравши функцію активації з простими похідними. Таким чином ми зменшуємо кількість арифметичних операцій на процесорі. Інші типи мереж мають надто складні структури, які потребують додаткових витрат процесорного часу для врахування топології. Тому в якості методу навчання обрано алгоритм зворотного поширення помилок, а в якості типу мережі – багатошаровий персептрон.

Однак використання лише одношарової мережі призводить до поганого розпізнавання та якості навчання. Оскільки однією з вимог системи є постійна адаптація та швидке реагування на зміни, необхідно використовувати багатошарову мережу [10].

Тому була обрана багатошарова мережа з алгоритмом навчання: зворотне поширення помилок для реалізації.

Для покращення результатів розпізнавання буде використано семантичний фільтр на основі спеціальної структури (бази даних), яка здатна розпізнавати ймовірність помилок введення.

2.3. Модифікація методу подання зображення до нейромережі

Існує кілька варіантів подачі зображення на вхід структури нейронної мережі. Керуючись критеріями швидкості, необхідно вибрати спосіб, який мінімізує розмір мережі.

Одним із варіантів обробки зображень є використання сіткової апроксимації. Для цього вихідне зображення розбивається на блоки. Далі кожен

блок отримує певне значення, зазвичай 0 або 1, де 1 означає, що блок містить частину символу, а 0 означає, що блок порожній. Приклад наведено на рис. 2.6.

Чим вище роздільна здатність сітки, тим точніші результати дає цей метод. Однак навіть при розмірі 16X16 на вході нейронної мережі знаходиться 256 пікселів, або 256 сигналів.

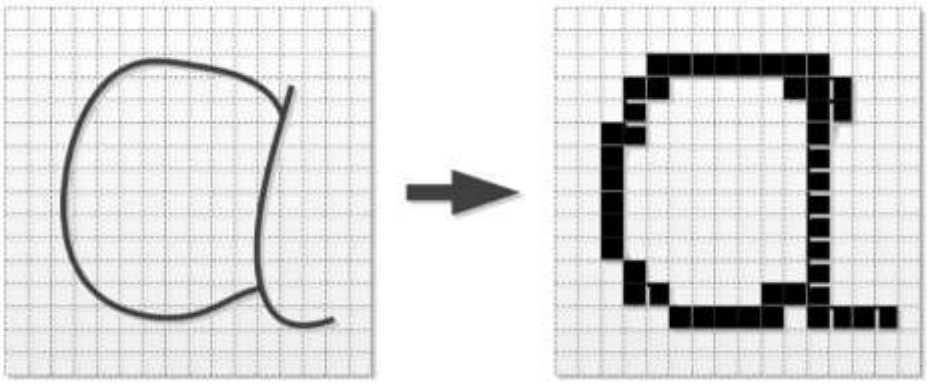


Рисунок 2.6 – Пікселізація зображення

Отже, такий підхід, очевидно, призводить до збільшення розміру мережі, тому слід розглянути інший підхід.

Другий варіант – використання векторизації зображення. У цьому випадку зображення символу перетворюється в еквівалентний вектор, а векторне сумарне перетворення подається в мережу. Приклад такого процесу показаний на рис. 2.7.

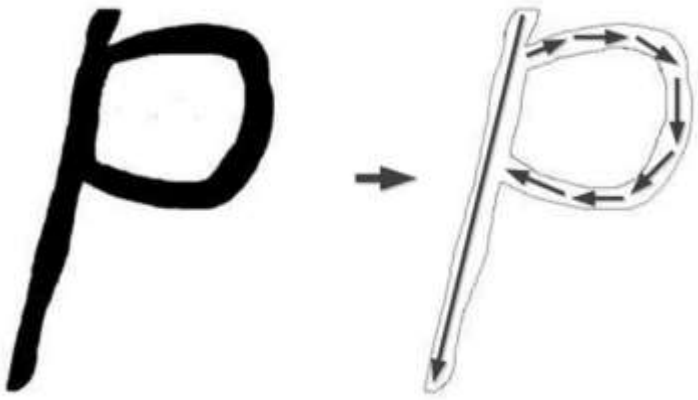


Рисунок 2.7 – Векторизація зображення

Такий підхід зменшує розмір мережі. Проте векторизація – досить складний процес, який вимагає тривалої обробки зображення. Тому в якості методу була обрана модифікація методу векторизації.

Для цього пікселізація буде виконана перед векторизацією, а векторизація буде виконана з використанням положення центру пікселя, максимально оптимізуючи кількість векторів. Такий підхід лише трохи збільшує кількість векторів порівняно з процесом гнучкої векторизації.

Але головне те, що, на відміну від пікселізації, для введення нейронної мережі не використовуються непотрібні блоки, тому розмір вхідного шару зберігається на належному рівні.

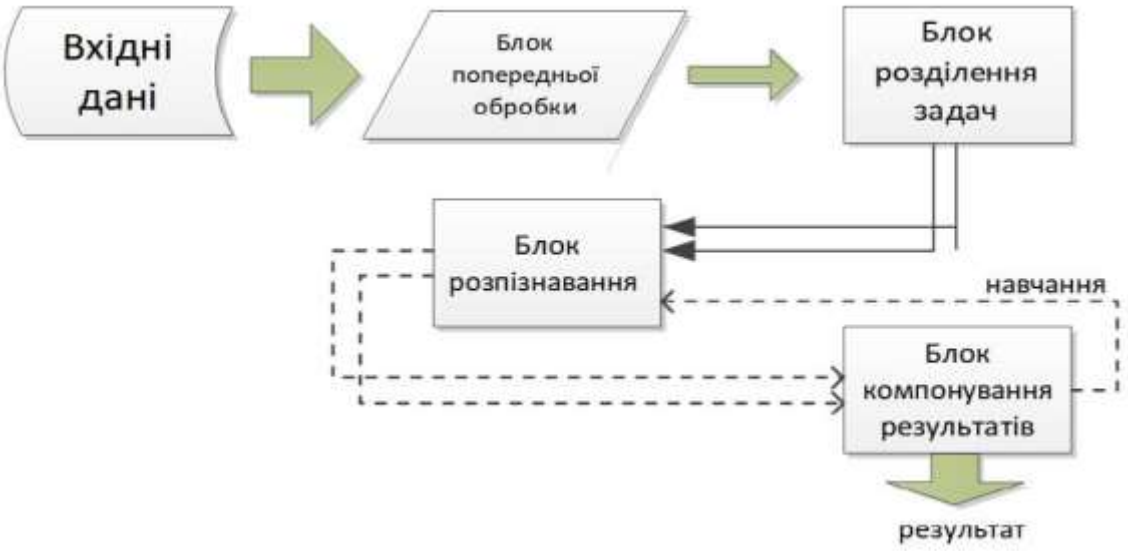
2.4 Архітектура та модель розпізнавання символів на тлі завад

Для успішного та швидкого впровадження інтелектуальних технологій необхідно розробляти інформаційні системи багаторівневим способом. На першому етапі ми розглянемо основні робочі компоненти та опишемо більш детально, як саме необхідні компоненти можуть бути реалізовані.

Розглядаючи блоковий поділ інформаційної системи, отримуємо її загальну архітектуру (рисунок 2.8)

Вхідні дані надходять безпосередньо в блок попередньої обробки. Цей блок розбиває зображення, вибирає символи та передає області на вхід блоку розділення завдань. Цього можна досягти за допомогою простого алгоритму обертання та вибору. Так, наприклад, достатньо знайти лінію поділу лінії і обробити лінію за досить простим принципом - поблочно виділяти символи/прогалини.

Блок розділення завдань вирішує, які символи належать до якого слова, і передає набори символів слово за словом, які далі обробляються нейронною мережею в блоки для розпізнавання. Блок розділення також повинен вирішити, які символи слід вважати спотвореними, надто шумними тощо.



Наступний блок – блок ідентифікації. Його завдання – визначити нечітку множину відповідності між зображенням і певним символом на основі заданого символу [17]. Блок розпізнавання повинен бути реалізований за допомогою нейронної мережі, що дозволить його перенавчати під час роботи. Ідентифікаційні дані передаються до наступного блоку, блоку синтезу результату, у формі заданого вектора.

Блок шукає в базі даних споріднені слова за знайденими словами, формуючи результати. Пошук реалізований на основі рекурсивного алгоритму. Для ефективної реалізації вам потрібно додати кешування та відсікти непотрібні рішення. Результати повертаються користувачеві у вигляді розпізнаних слів. Однак блоки компонування виконують іншу мету.

Після розпізнавання слова блок передає цю інформацію блоку розпізнавання, щоб той, у свою чергу, перевірив себе та виправив свої помилки. Тут дуже важливо, щоб блок розпізнавання був реалізований у вигляді нейронної мережі з усіма необхідними властивостями.

Для навчання мережі буде використовуватися алгоритм зворотного поширення.

Тобто, визначити похибку обробки:

$$E(w) = \frac{1}{2}\sum_{j,p}(y_{j,p}^{(N)} - d_{j,p})^2 \qquad (2.33)$$

На основі якої виконується градієнтний спуск для налаштування ваг мережі:

$$\Delta w_{ij}^{(n)} = -\mu \frac{\partial E}{\partial w_{ij}} \tag{2.33}$$

Для внутрішнього шару використовується обрахунок аналогічним чином:

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial y_j} \frac{dy_j}{ds_j} \frac{\partial s_j}{\partial w_{ij}} \tag{2.34}$$

Очевидно, вам потрібно вибрати функцію активації.

Наступні типи сигмоїди найчастіше використовуються як функції активації:

- функція Фермі (експоненціальна сигмоїда),
- раціональна сигма,
- Гіперболічний тангенс.

Час обробки для обчислення раціональної сигмоїдної функції активації є найкоротшим з усіх сигмоїд. Обчислення гіперболічного тангенса займає найбільше процесорних циклів. Сигмоїда дуже повільно обчислюється порівняно з функцією активації порогу. Якщо ви можете розпочати порівняння з деяким значенням (порогом) відразу після підсумовування в пороговій функції, то у випадку сигмоїдної функції активації вам потрібно обчислити сигмоїд (у кращому випадку для виконання трьох операцій потрібен час: за модулем, додавання та ділення) перед порівнянням із порогом (наприклад, нуль). Якщо врахувати, що всі прості операції обчислюються процесором приблизно за однаковий проміжок часу, то робота сумованої сигмоїдальної функції активації (яка займе стільки ж часу) буде повільнішою за порогову функцію активації на пропорції 1:4.

Враховуючи вищесказане, розрахунок буде виконано за допомогою функції Фермі, оскільки вона має просту похідну, її легко обчислити та підходить для ідентифікації нейронів.

Розглянемо підсистему розпізнавання слів.

Як зазначалося вище, у цьому випадку програма розпізнавання символів повинна повернути вектор нечітких наборів [17] для слів.

Вектор нечіткої множини подібності символів – це впорядкований набір нормалізованих нечітких множин, що представляє ймовірність того, що зображення відповідає символу в наборі всіх символів:

$$v = \begin{matrix} < \{(a: p_{1,1}), (b: p_{1,2}), \dots, (z: p_{1,n})\}, \\ (a: p_{2,1}), (b: p_{2,2}), \dots, (z: p_{2,n}) \\ \vdots \\ (a: p_{m,1}), (b: p_{m,2}), \dots, (z: p_{m,n}) > \end{matrix} \quad (2.35)$$

де a,b,…,z – символи алфавіту, а px,i – відповідно ймовірності відповідності x-вого зображення символа i-тому символу алфавіту.

Знання про відповідність між векторами та словами, отримані з бази даних, далі використовуються для визначення результатів розпізнавання та подальшого вдосконалення системи оптичного розпізнавання.

Давайте детальніше розглянемо слово база даних. Ця база може повністю складатися з найбільш часто вживаних слів.

Вектор нечіткої множини отримується з підсистеми керування [17]. Для кожного слова в базі даних ми обчислимо ймовірність того, що даний вектор відповідає певному слову.

Припустимо, що відсутність або наявність деяких додаткових символів погіршить ймовірність у λ разів. потім можна обчислити за рекурсивною формулою:

$$f(s, v) = \max(f(s.tail, v.tail) \cdot v.head[s.head], \lambda f(s.tail, v), \lambda(s, v.tail)) \quad (2.36)$$

$$f(s, []) = \lambda f(s.tail, []) \quad (2.37)$$

$$f([], v) = \lambda f([], v.tail) \quad (2.38)$$

$$f([], []) = 1 \quad (2.39)$$

де f(s,v) – ймовірність, що слово s відповідає вектору v. X.head – перший символ вектора/слова, а X.tail – вектор/слово без першого символу. [] –

відповідно позначає пустий вектор/стрічку. v.head[x] – описує ймовірність першого символа вектора відповідати деякому символу x.

Як бачите, останні три рівняння є основою для рекурсії.

Ви можете скористатися простішою обробкою, просто видаливши префікси та суфікси або перевіривши слова. Усі ці способи можуть прискорити роботу алгоритму, але можуть некоректно чи неточно вирішити задачу визначення відповідностей.

Основні 100 000 найбільш вживаних слів англійської мови. Взагалі стільки слів зайві. Тому необхідно розробити методи оптимізації, тобто розбити слова на вживані групи і спочатку перевірити збіги з найбільш часто вживаними словами, а потім переходити на інший рівень. Наприклад, якщо на першому рівні знайдено слово з рейтингом вище певного порогу, пошук можна зупинити.

Цю оптимізацію можна покращити, створивши графіки подібності. Якщо слово має високу ймовірність відповідності, тоді ви можете розглянути кілька схожих слів, можливо, вони матимуть вищу ймовірність відповідності.

Такий граф потрібно побудувати лише один раз, тому складність побудови нас зовсім не цікавить. Ви можете використовувати просту побудову графіка на основі відстані Левенштейна. А поріг подібності дуже низький, наприклад різниця в 1-2 символи.

Зібравши разом усе, що було сказано в цьому розділі, ми маємо вже досяжну систему, яка буде виконувати щось подібне до асоціативного пошуку в базі даних.

Загалом слід зазначити, що алгоритм подібності є рекурсивним і повільним у виконанні, при простому аналізі швидкість його роботи може досягати O(3N). Для оптимізації рекомендується використовувати кеш результатів алгоритму, щоб пошук можна було зберегти на позначці O(N2). Іншим прискоренням є використання алгоритму відсікання на основі результатів. Це не покращить оцінки найгіршого випадку, але значно пришвидшить роботу в середньому випадку.

2.5 Складові додатку розпізнавання символів на тлі завад

Інформаційна технологія розпізнавання символів у контексті інтерференції з нейронними мережами на основі зворотного поширення помилок має дві складові: навчання та безпосередню роботу нейронної мережі (рис. 2.9). Сформулюємо фази цих компонентів.

Навчання (проводиться при завантаженні файлу з еталонними зображеннями символів для розпізнавання (шрифту)), складається з наступних етапів:

1. Процес створення нейронної мережі, яка повертає помилки.

2. Ітераційний процес обчислення значень вагових коефіцієнтів нейронів у кожному шарі мережі (пошук вагової матриці W кожного шару).



Рисунок 2.9 – Структура додатку розпізнавання символів на тлі завад

Функціонування:

1. Процес поділу зображення на блоки символів.

2. Процес пікселізації непорожніх блоків.

3. Процес блокової векторизації.

4. Процес подачі вектора вхідного сигналу x на вхід мережі.

5. Процес розрахунку вихідних сигналів нейронів кожного шару нейронної мережі.

6. Процес визначення нечіткого набору, що зображення відповідає символу.

7. Процес об'єднання результатів розпізнавання в окремі слова.

8. Процес виправлення помилок ідентифікації.

У цій частині роботи здійснюється теоретична демонстрація розвитку інформаційної технології розпізнавання символів на фоні перешкод.

Теоретичною основою реалізації інформаційних технологій доцільно обрати штучну нейронну мережу, а нейронну мережу зворотного поширення помилок – як мережу, що забезпечує найвищу надійність і швидкість розпізнавання символів в умовах перешкод. Проаналізовано структуру та математичну модель нейронної мережі зворотного розповсюдження помилок, яка з метою підвищення надійності розпізнавання символів на тлі перешкод модифікована таким чином: Метод зворотного поширення помилок у навчанні нейронної мережі розраховано за допомогою функціх Фермі. Активація нейронів покращує метод, який покращує векторизацію зображень шляхом пікселізації перед векторизацією, що може спростити структуру нейронної мережі. Розроблено процес обробки даних для запропонованої інформаційної технології та його компоненти: послідовність навчання та операцій.

РОЗДІЛ 3 РЕАЛІЗАЦІЯ ЗАСТОСУНКУ ДЛЯ РОЗПІЗНАВАННЯ СИМВОЛІВ НА ТЛІ ЗАВАД НА ОСНОВІ НЕЙРОМЕРЕЖІ

3.1. Розробка алгоритму безпосередньо розпізнавання символів на тлі завад

Загальний алгоритм роботи програми розпізнавання символів на фоні перешкод наведено на рисунку 3.1.

Для коректного та швидкого проектування використовуються парадигми об'єктно-орієнтованого програмування та можливість автоматичного проектування систем за допомогою UML. Розроблена архітектура класу повністю відображає архітектуру програми, описану в розділі 2.4.

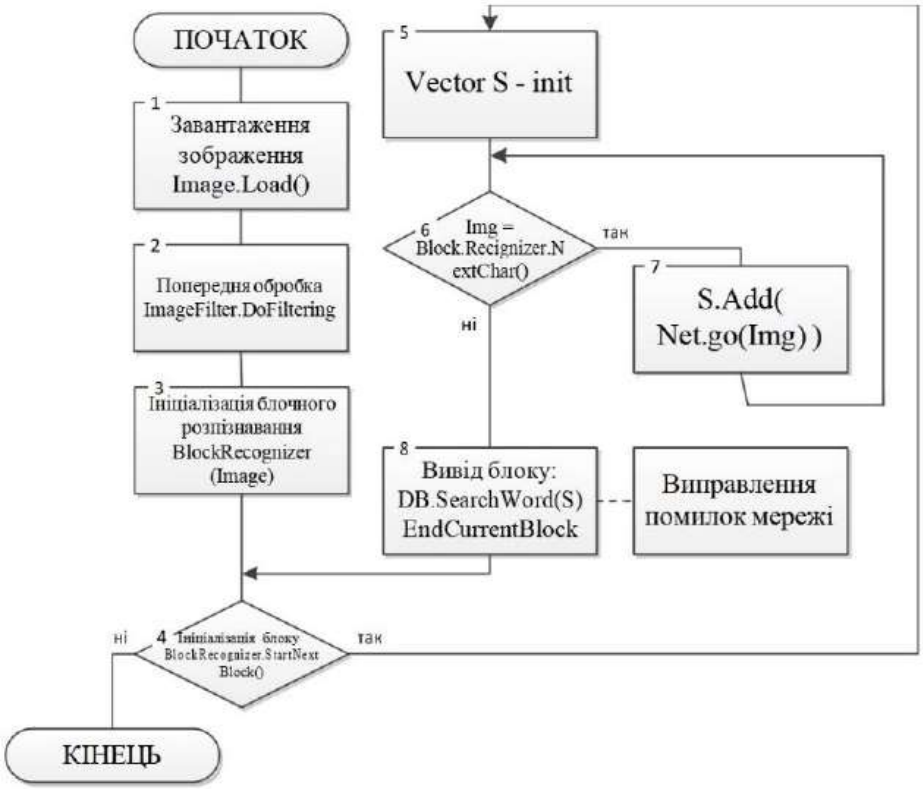


Рисунок 3.1 – Алгоритм безпосередньо розпізнавання символів на тлі завад

Програмне забезпечення складається з кількох модулів (рис. 3.2). Модуль NeuNet-Layer-Vertex виконує оптичне розпізнавання.

Загалом, три класи реалізують нейронну мережу. Навчання нейронної мережі реалізовано на основі алгоритму зворотного поширення. Нейронна мережа має трирівневу структуру:

1. Вхідний – має розмір 256 – 1024 нейронів;

2. Внутрішні - розміром 200-500 нейронів;

3. Оригінальний розмір 26.

Іншим важливим компонентом є клас Searcher - він реалізує роботу над базою слів і виконує необхідні алгоритми для пошуку підходящих слів. Реалізовано на основі динамічного програмування. Таке впровадження скорочує витрати часу.

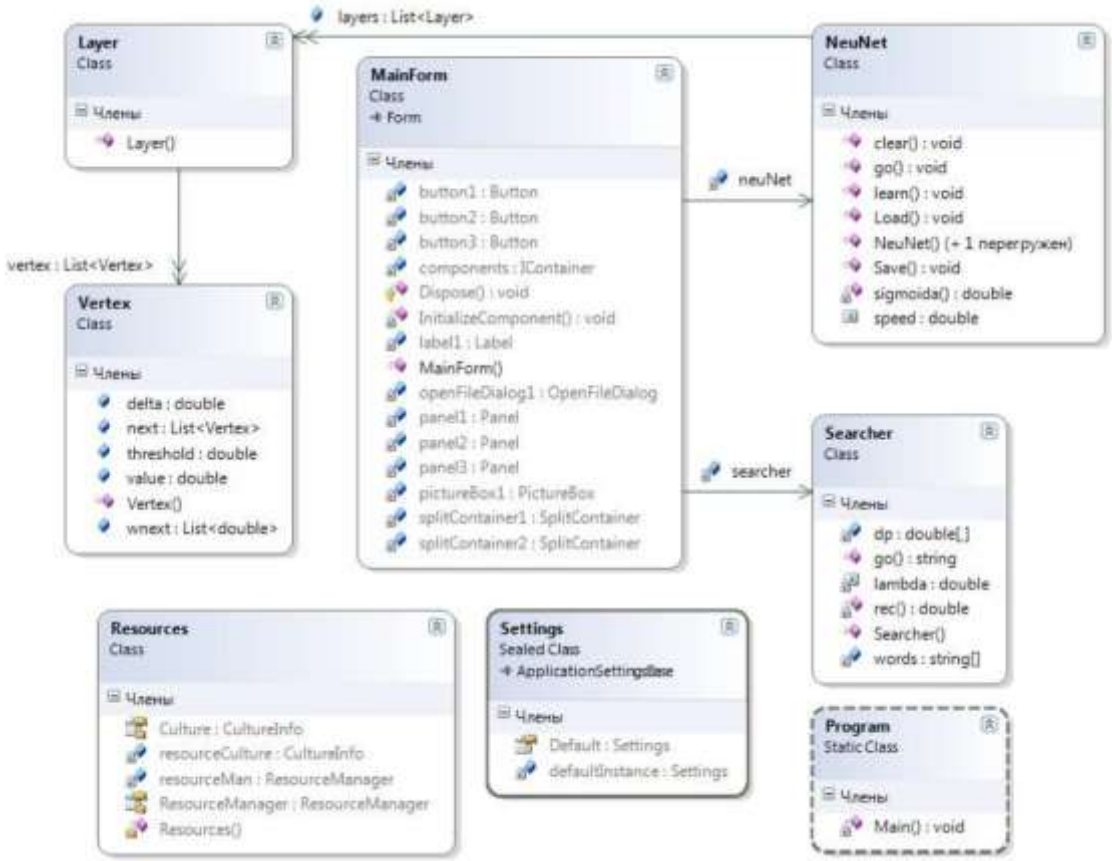


Рисунок 3.2 – Діаграма класів

Алгоритм пошуку: function F(S,V,a,b)

Вхідні дані: S – слово, V – вектор, a – поточна позиція в слові, b- поточна позиція в векторі;

Вихідні дані: Ймовірність відповідності вектора слову;

1. Якщо a вказує на кінець слова AND b вказує на кінець вектора – повернути 1.0 інакше перейти до кроку 2;

2. Якщо a вказує на кінець слова - повернути λF(S,V,a,b+1) інакше перейти до кроку 3;

3. Якщо b вказує на кінець вектора – повернути λF(S,V,a+1,b) інакше перейти до кроку 4;

4. Якщо функція вже була визвана з цими параметрами – повернути записане в кеш значення результату, інакше перейти до кроку 5;

5. Обрати як результат максимальне з трьох значеннь: a. F(S,V,a+1,b+1) * V[b][S[a]];

b. λF(S,V,a+1,b);

c. λF(S,V,a,b+1);

Записати значення до кешу та повернути як результат.

Інші класи використовуються як допоміжні для розробки графічних інтерфейсів. Основний клас форми виконує функції попередньої обробки та призначення завдань. Також прийнято рішення щодо продовження онлайн-навчання.

Для розробки цього застосунку була обрана мова програмування C#, яка скомпільована з MIL. Це дозволяє програмам працювати в будь-якому середовищі, де встановлено віртуальну машину .NET. Середовищем розробки обрано пакет Microsoft Visual Studio. Це пояснюється тим, що середовище розробки Visual Studio є дуже поширеним і простим у використанні. Безкоштовна версія "урізаної" версії сімейства Visual Studio Express також вступає в дію.

Microsoft Visual C# [18, 19] – це інтегроване середовище розробки програм для мови C#, розроблене Microsoft і доступне окремо як частина набору Microsoft Visual Studio або як безкоштовний пакет Visual C# Express Edition з обмеженим набором функцій [19].

3.2. Програні елементи розпізнавання символів на тлі завад

Програма складається з двох збірок: iTell.exe і iTellUtils.dll. У першій збірці, крім інтерфейсу редактора для налаштування виділення блоку символів, присутній графічний інтерфейс користувача.

Збірка iTell містить такі основні класи:

- Console і ConsoleForm - класи, які інкапсулюють роботу консолі та графічний інтерфейс консолі відповідно;

- DBForm - редактор тезаурусів;

- LearningForm - редактор для роботи з нейронними мережами;

- MainForm - головне вікно, що містить інтерфейс користувача і контролює всі інші процеси комплексу;

- Програма - стандартний клас для точок входу програми;

- Recognizer - помічник, який з'єднує інтерфейс користувача з системою розпізнавання;

- SettingsForm і UserSettings - вікно налаштувань користувача та клас, який інкапсулює налаштування користувача.

Збірку iTellUtils можна використовувати незалежно від системи.

Ця збірка містить клас для роботи з нейронними мережами: NeuralNetwork. Це дозволяє використовувати структуру мережі, навчати її та запускати на вхідних даних. Цей клас може використовувати будь-яку кількість шарів. Для запису файлів використовується звичайний текстовий формат (для можливості ручного редагування).

Для виконання обчислень через мережу на вхід подається вектор значень від рецепторів - дробові значення. Після запуску функції Run() ви можете використовувати функцію Output(), щоб отримати значення на виході нейронної мережі. Для навчання функція Learn(a,b) використовується з двома аргументами: перший

є вхідними даними, а другий – вихідним рівнем, необхідним для мережі.

Клас WordDB інкапсулює роботу тезауруса. Цей клас дозволяє створювати порожню базу даних, завантажувати базу даних із файлу (формат файлу – це список слів, по одному на стрічку, виберіть простий формат для легкого ручного редагування). Цей клас містить функціональні можливості для збереження бази даних у файлі, що дозволяє додавати нові слова, виконувати

пошукові запити за префіксами. Однак основна функція – знайти відповідне слово.

Це виконується функцією findBestWords(iTellUtils.FuzzyVector), яка отримує вектор нечіткого набору як параметр і таким чином створює масив слів, які найкраще відповідають вектору нечіткого набору.

Іншим важливим класом є клас ImageFilter, який можна використовувати для попередньої обробки зображень. Цей клас виконує фільтрацію на основі фільтрів ранжирування.

Клас BlockRecognizer ділить зображення на блоки. Функція StartImage використовується для завантаження зображення. Потім виконується функція Next*Char, доки не буде повернено порожній об’єкт, що вказує на завершення зображення. Основні класи показані на рисунку. 3.3.

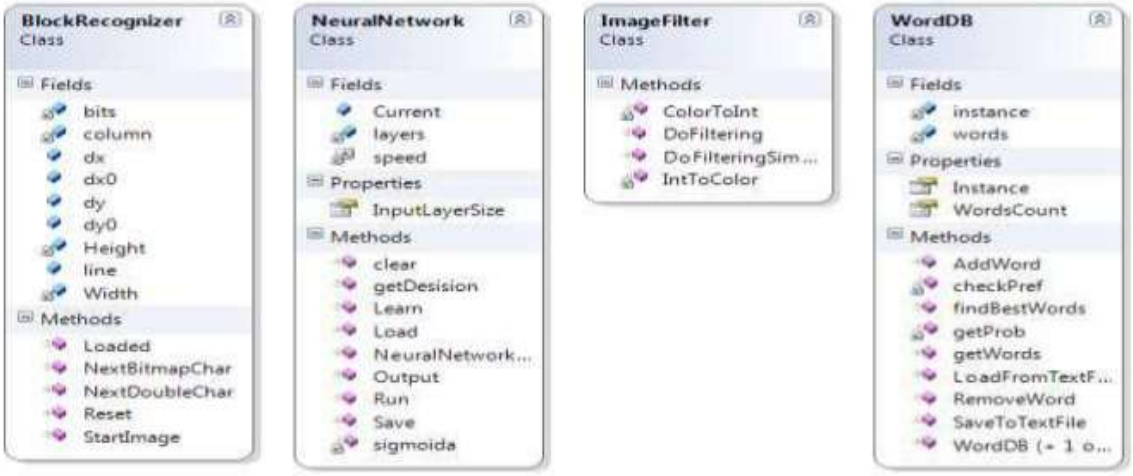


Рисунок 3.3 – Основні класи програми

Одразу після запуску програми з’являється головне вікно (рис. 3.4). Головне вікно повністю відповідає стандартному інтерфейсу Windows. Вікно складається з робочої області та віконного меню.

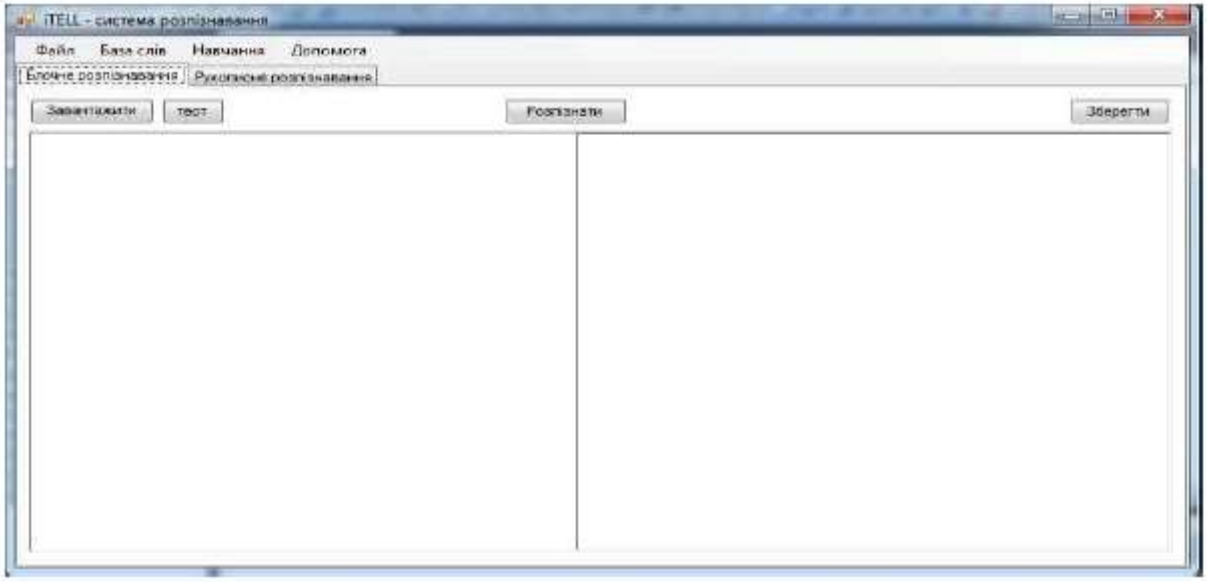


Рисунок 3.4 – Головне вікно програми

Для ідентифікації достатньо кількох кліків. Для цього необхідно завантажити зображення (за допомогою кнопки "Завантажити" або відповідного підпункту меню файлів). Далі виконайте розпізнавання ("Розпізнавання") і результати відобразяться в полі виведення (права частина вікна). Приклад результату розпізнавання наведено на рисунку 3.5.

Програма складається з модуля налаштування зображення, модуля тезауруса, модуля навчання та генерації нейронної мережі та інших модулів.

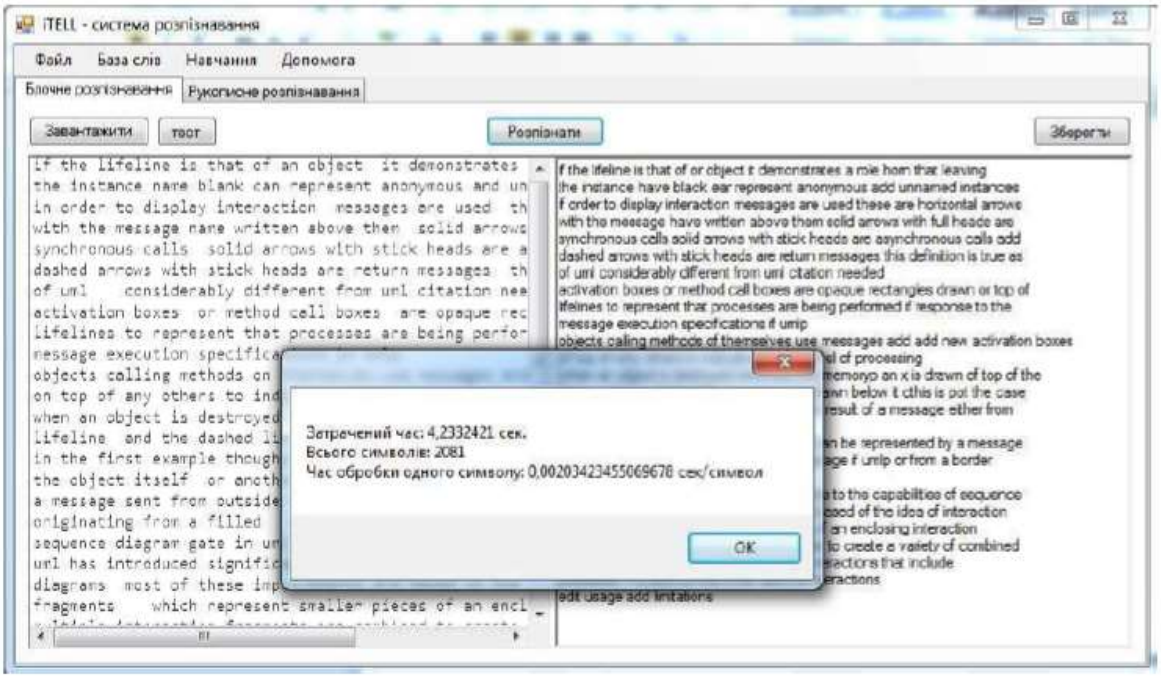


Рисунок 3.5 – Результат розпізнавання

По-перше, врахуйте налаштування користувача. Використовуйте пункт меню "Файл > Параметри", щоб викликати налаштування. Вікно налаштувань показано на малюнку 3.6.

Ці налаштування дозволяють вибрати файл нейронної мережі та файл бази слів за замовчуванням. Для редагування бази слів скористайтеся редактором "База слів > Редагувати базу". Редактор (рис. 3.7) дозволяє додавати слова, видаляти групи слів і здійснювати пошук у базі даних. Програма дозволяє зберігати бази слів у файли у звичайному форматі.

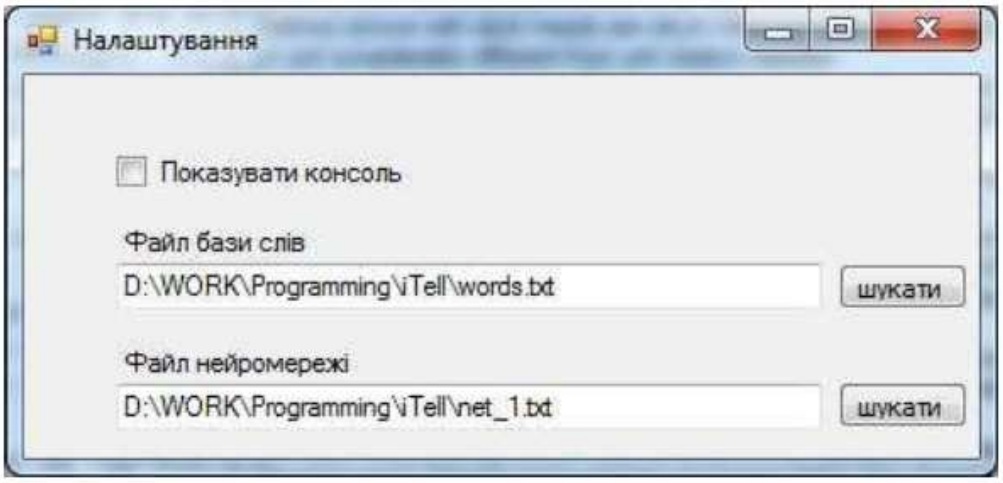


Рисунок 3.6 – Вікно налаштувань

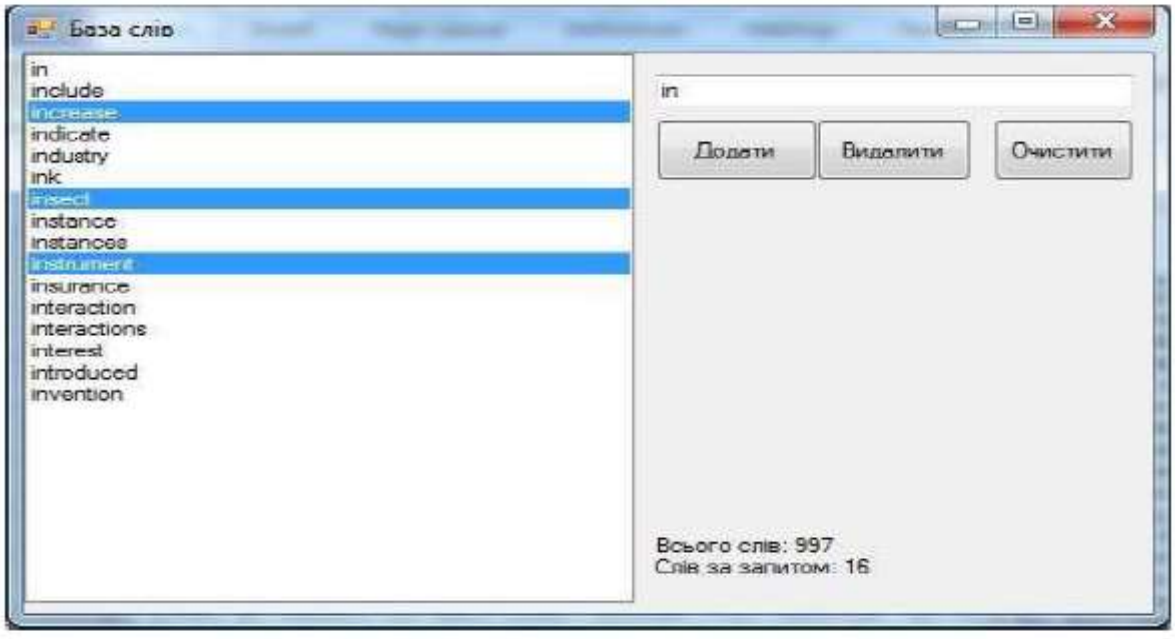


Рисунок 3.7 – Редактор бази слів

Для вивчення/завантаження/збереження нейронних мереж ми розробили редактор, доступний у розділі "Навчання". Для навчання потрібно вибрати каталог із прикладами символів (кожен символ в окремому файлі, формат зображення BMP). Далі виберіть кількість ітерацій, які потрібно виконати на наборі, і натисніть "Почати навчання". Коли мережа тренується, деякі статистичні дані будуть відображатися в полі виводу внизу вікна (рис. 3.8).

У цьому полі відображається кількість правильно визначених спроб і загальна кількість спроб. Для покращення видимості відображається стрічка,

на якій кожна спроба в хронологічному порядку замінюється символом "+" або "o", що вказує на те, що спроба була ідентифікована правильно чи неправильно відповідно.

Для точного налаштування модуля обробки блоків можна викликати додатковий редактор з головного вікна, натиснувши кнопку "Тестувати" праворуч від кнопки "Завантажити".

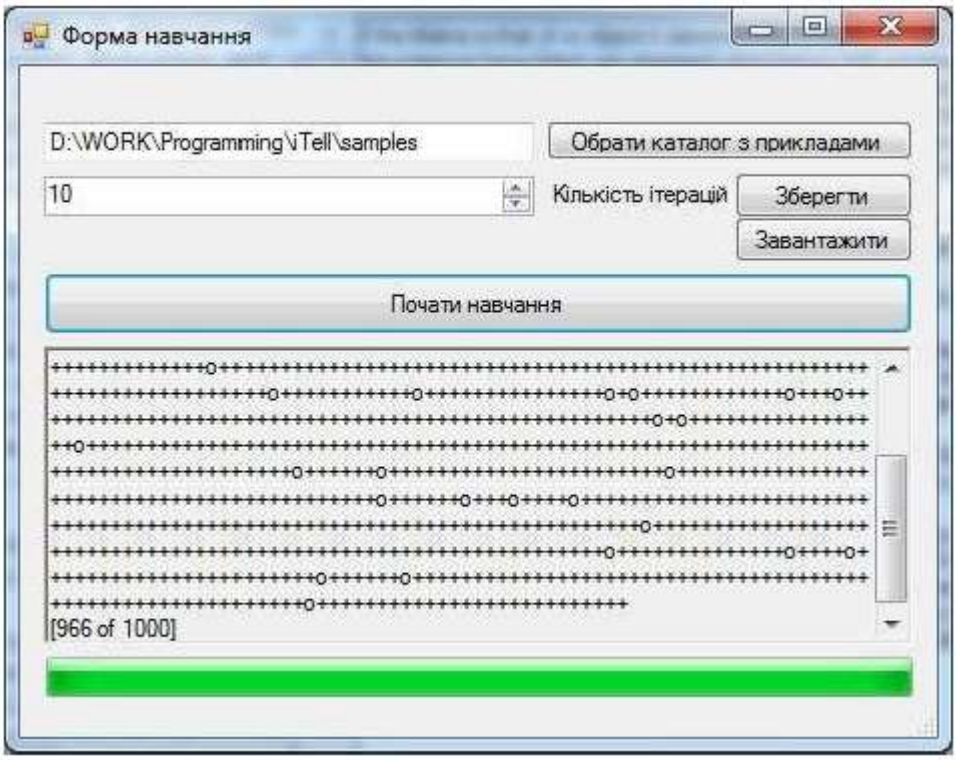


Рисунок 3.8 – Навчання нейронної мережі

Редактор зображено на рисунку 3.9. Це дозволяє вам точно налаштувати обробку блоків, наприклад додавати зміщення або встановлювати розміри символів. Можна також використовувати для підготовки символів для навчання мережі.

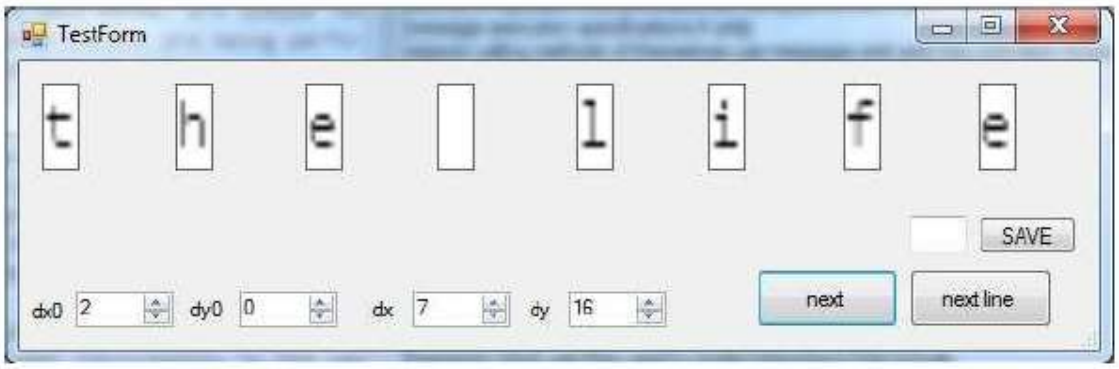


Рисунок 3.9 – Налаштування модуля блочної обробки

3.3. Тестування та аналіз результатів роботи додатку розпізнавання символів на тлі завад

Для аналізу роботи програми було використано навчання нейронної мережі з 10 000 випадкових символів. Візьмемо 1000 найпоширеніших слів англійською як тезаурус і використаємо статтю Вікіпедії як текст для ідентифікації.

Перший тест – це чисте зображення без відволікаючих факторів.

З початку розпізнавання (рис. 3.10) програма розпізнала майже всі символи з точністю 98,5%.

Швидкість обробки 2107 символів становить 1,7 секунди.

Середній час обробки становив 0,0008 секунди. Така швидкість пояснюється використанням простих нейронних мереж і перенесенням обчислювальної складності на блок обробки результатів.

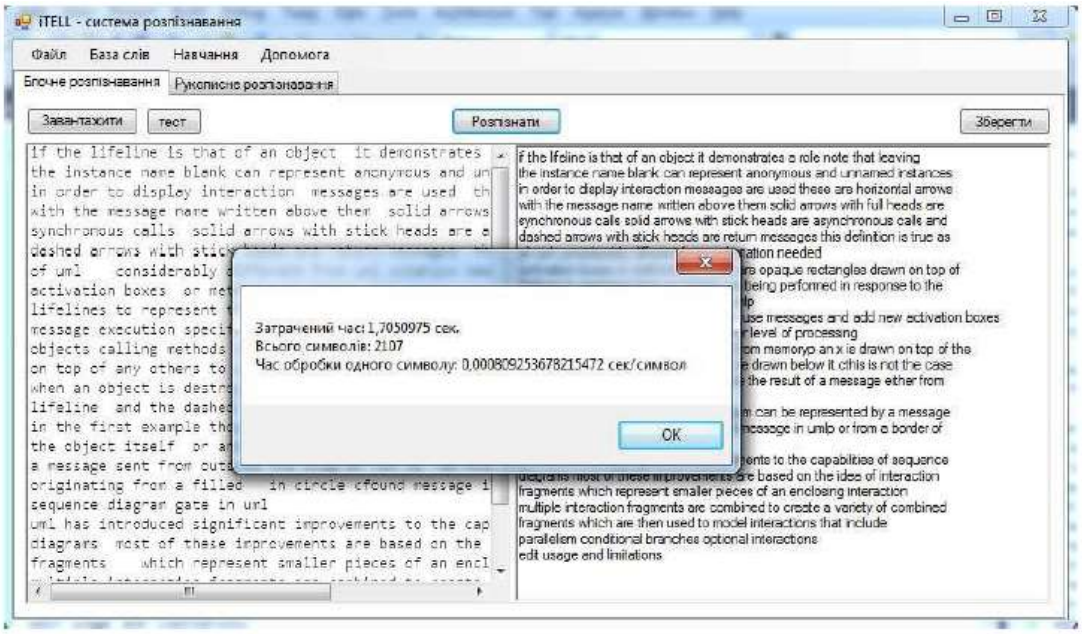


Рисунок 3.10 – Тестове розпізнавання зображення без завад

Були отримані наступні результати у вигляді розмиття та шуму для 5% порушення: Час обробки: 1,6 секунди Точність: 92,4% (залишилося в межах встановленої цілі). Подальше збільшення перешкод призвело до таких результатів:

10% - Показник точності становить 82,3%, а швидкість залишається незмінною;

20% - Показник точності становить 70,4%, а швидкість залишається незмінною;

Пошкоджений текст (випадковий нерегулярний шум, вставка додаткових чисел, неправильні символи під час відображення) - точність 85,971% постійної швидкості.

Для визначення сильних сторін розробленої програми було порівняно надійність розпізнавання символів у контексті інтерференції між розробленою та змодельованою програмами. Як програму моделювання було обрано програму MeOCR 1.0 [14]. Він перетворює скановані документи на текстові документи, які можна редагувати, і експортує їх у Microsoft Word. Результати дослідження надійності розпізнавання двома програмами одного тексту з

різними значеннями шуму наведені в таблиці. 3.1. Приклад тексту з перешкодами показано на малюнку. 3.11.

Таблиця 3.1 - Достовірність розпізнавання символів на тлі завад

	Наявність завад			
	0%	5%	10%	20%
Розроблена програма	98,5%	92,4%	82,3%	70,4%,
Аналог (програма MeOCR)	98,1%	90,5%	80,2%	67,9%,
Різниця	0,4%	1,9%	2,1%	2,5%

Огляд результатів тестування (табл. 3.1) показує, що розроблена програма має кращий показник надійності розпізнавання символів на фоні перешкод. Таким чином, поставлені цілі були досягнуті, оскільки надійність

ідентифікації розробленої програми вища, ніж у симуляції (0% шуму - на 0,4% вище, 5% шуму - на 1,9% вище, 10% шуму - на 2,1% вище, 20% шум - на 2,5% вище). Для чистих зображень якість розпізнавання наближається до 100%. На практиці частота відмов становить менше 1%, а за такими вхідними даними середня похибка комплексу не перевищує 2,2%.

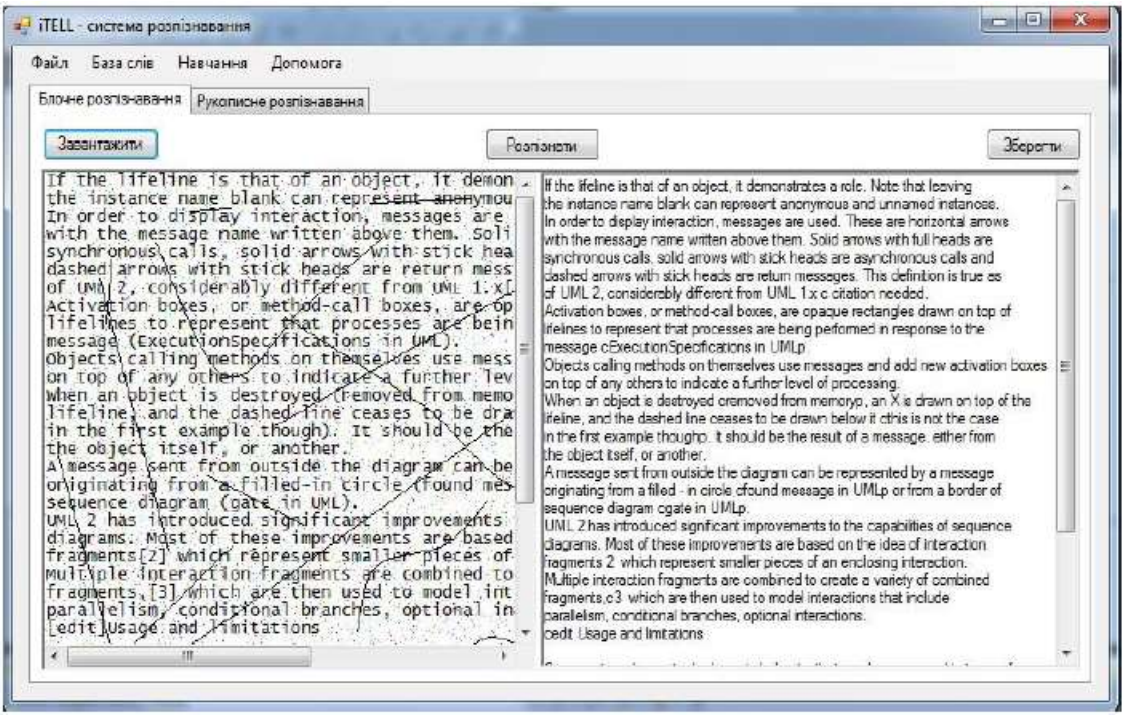


Рисунок 3.11 – Тестове розпізнавання зображення текста з завадами

Програма реалізована на базі платформи .NET. Хоча платформа .NET не є найефективнішим рішенням з точки зору продуктивності, вона допомагає перенести складність програмного забезпечення на інші платформи.

У цьому розділі реалізувалися алгоритми роботи програмних засобів розпізнавання символів в умовах. Розроблений застосунок був реалізований на мові програмування Microsoft Visual C# і в середовищі програмування Microsoft Visual Studio. Також була проведена розробка програмного засобу для розпізнавання символів у контексті перешкод. Програма складається з двох збірок: iTell.exe і iTellUtils.dll. Тести показують надійну роботу розроблених програм. Аналіз результатів роботи доводить, що розроблена програма має кращий показник надійності розпізнавання символів, ніж змодельована програма на фоні перешкод.

ВИСНОВКИ

У кваліфікаційній роботі магістра розробка цієї інформаційної технології та програми розпізнавання знаків у контексті перешкод виявилася актуальною та вигідно порівняно з розробкою аналогів. Результати аналізу предметної області, яка враховує сучасні методи розпізнавання образів, сучасні методи попередньої обробки зображень, методи представлення зображень у нейронних мережах, а також основну топологію нейронних мереж і алгоритми їх навчання, які стали можливими. згідно з алгоритмом зворотного поширення помилок, щоб обґрунтувати вибір нейронної мережі багатошарового персептрона зі здатністю до навчання. Показано основні недоліки реалізації рішень на основі інших топологій і алгоритмів.

Всі поставлені завдання виконано. Також описано математичний апарат для виконання поставленого завдання, а саме розпізнавання символів у контексті перешкод. Наведено алгоритм роботи програми розпізнавання символів на фоні перешкод. Основною перевагою розробленої програми розпізнавання є нечіткий пошук очікуваного результату. Такий додатковий процес дозволяє, по-перше, більш якісно провести ідентифікацію, а по-друге, дозволяє ідентифікувати блоки символів з перешкодами.

Програмне забезпечення розроблено на мові програмування C# та скомпільовано для платформи .NET у середовищі розробки Microsoft Visual Studio. Цілі, поставлені в роботі, досягнуті, оскільки надійність розпізнавання розробленої програми вища, ніж у симуляції (0% шуму - на 0,4% вище, 5% шуму - на 1,9% вище, 10% шуму - на 2,1% вище , 20% шуму – 2,5% вище).

СПИСОК ЛІТЕРАТУРИ

1. Чуба Б. О. Програмний засіб розпізнавання символів на тлі завад на основі нейронної мережі / Б. О. Чуба, О. К. Колесницький // Науково-практична конференція «Сучасні тенденції розвитку системного програмування» (25-26 листопада 2016 р., Україна, м. Київ.
2. Алексєєв А. Алгоритм розпізнавання символів на основі структурного підходу / А. Алексєєв, В. Заяць, Д. Іванов // Вісник Нац. ун-ту „Львівська політехніка „Комп‘ютерна інженерія та інформаційні технології". – 2002. – № 468. – С.129 - 133.
3. Реалізація нейронних мереж [Електронний ресурс] // Режим доступу: http://alexandria.nu/ai/neural_net_demo
4. ABBYY FineReader [Електронний ресурс] // Режим доступу:http://help.abbyy.com/FineReader/FineReader12/Russian/QuickTasks/general.htm)
5. Програма CuneiForm [Електронний ресурс] // Режим доступу:http://programdownloadfree.com/load/text/scanning_recognition/ocr_cuneiform/74-1-0-131
6. MeOCR - OCR Image to Text Converter [Електронний ресурс] // Режим доступу: http://www.meocr.com/
7. Горбань А.Н. Методы Нейроинформатики, сборник научныхтрудов/ А.Н. Горбань. – Красноярск: КГТУ, 1998. – 341с.
8. Заяць В.М. Архітектура подіє – орієнтованих систем на прикладі системи розпізнавання рукописного тексту / В.М. Заяць, Д.О. Іванов //Вісник Нац. ун-ту „Львівська політехніка „Комп‘ютерна інженерія та інформаційні технології;. – 2004. – № 530. – С. 78 - 83.
9. ctypes — A foreign function library for Python – https://docs.python.org/3/library/ctypes.html
10. Browser & Platform Market Share: September 2020 – https://www.w3counter.com/globalstats.php
11. What are extensions? – https://developer.chrome.com/extensions

12. Densely Connected Convolutional Networks, Gao Huang, Zhuang Liu, Laurens van der Maaten, 28 Jan 2018 - https://arxiv.org/pdf/1608.06993.pdf
13. CSPNET: A NEW BACKBONE THAT CAN ENHANCE LEARNING CAPABILITY OF CNN, Chien-Yao Wang, Hong-Yuan Mark Liao, I-Hau Yeh, Yueh-Hua Wu, Ping-Yang Chen, Jun-Wei Hsieh, 27 Nov 2019
14. YOLOv3: An Incremental Improvement, Joseph Redmon, Ali Farhadi, 8 Apr 2018
15. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition, Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, 23 Apr 2015
16. DC-SPP-YOLO: Dense Connection and Spatial Pyramid Pooling Based YOLO for Object Detection, Zhanchao Huang, Jianlin Wang, College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China
17. CBAM: Convolutional Block Attention Module, Sanghyun Woo, Jongchan Park, Joon-Young Lee, and In So Kweon, 18 Jul 2018
18. YOLOv4: Optimal Speed and Accuracy of Object Detection, Alexey Bochkovskiy, Chien-Yao Wang, Hong-Yuan Mark Liao, 23 Apr 2020
19. SEARCHING FOR ACTIVATION FUNCTIONS, Prajit Ramachandran, Barret Zoph, Quoc V. Le, 27 Oct 2017
20. Mish: A Self Regularized Non-Monotonic Activation Function, Diganta Misra, Landskape KIIT, Bhubaneswar, India, 13 Aug 2020
21. CUDA Toolkit – https://developer.nvidia.com/cuda-toolkit
22. NVIDIA cuDNN – https://developer.nvidia.com/cudnn
23. OpenCV: Introduction – https://docs.opencv.org/master/d1/dfb/intro.html
24. Yolo v4, v3 and v2 for Windows and Linux – https://github.com/AlexeyAB/darknet
25. AlexeyAB/darknet/src/detector.c – https://github.com/AlexeyAB/darknet/blob/d65909fbea471d06e52a2e4a4113280dc2edaa6/src/detector.c#L1640

26. Base64 – https://developer.mozilla.org/en-US/docs/Glossary/Base64
27. What Socket.IO is – https://socket.io/docs/
28. "The AV-TEST Security Report", 2019/2020.: https://www.av-test.org/fileadmin/pdf/security_report/AV-TEST_Security_Report_2019-2020.pdf., AV-TEST, 2019 2. X. Wang, D. Xuan, X. Fu, A. Champion, and W. Yu, "Detecting worms via mining dynamic program execution", Proc. 3rd, SecureComm, 2007.
29. M. Karresand, "Separating Trojan horses, viruses, and worms - A proposed taxonomy of software weapons", IEEE, 2003
30. N. Scaife, K. R. B. Butler, P. Traynor, and H. Carter, "CryptoLock (and Drop It): Stopping Ransomware Attacks on User Data", Proc. – Int. Conf. Dist. Comp. Sys., 2016.
31. A. Zaki and B. Humphrey, "Unveiling the kernel : Rootkit discovery using selective automated kernel memory differencing", Virus Bulletin, 2014.
32. Y. Ye, S. S. Iyengar, D. Adjeroh, and T. Li, "A Survey on Malware Detection Using Data Mining Techniques", ACM Comp. Survey, 2017.
33. N. B. Ithnin and G. A. N. Mohamed, "Survey on Representation Techniques for Malware Detection System", American Journal of Applied Science, 2017.
34. A. Damodaran, T. H. Austin, F. C. A. Visaggio, Di Troia, and M. Stamp, "A comparison of static, dynamic, and hybrid analysis for malware detection", J. Comp. Vir. Hack. Tech., 2017.
35. M. Chowdhury and A. Rahman, "Malware Analysis and Detection Using Data Mining and Machine Learning Classification", 2018.
36. Z. Bazrafshan, S. M. H. Fard, H. Hashemi, and A. Hamzeh, "A survey on heuristic malware detection techniques", IKT 2013.
37. I. You and K. Yim, "Malware obfuscation techniques: A brief survey", BWCCA 2010.
38. R. Hosseini and A. Souri, "A state-of-the-art survey of malware detection approaches using data mining techniques", Hum. Comp. Inf. Scie., 2018.

39. M. Alazab, P. Watters, and S. Venkataraman, "Towards understanding malware behaviour by the extraction of API calls", CTC 2010.
40. W. Wong and M. Stamp, "Hunting for metamorphic engines", Journal of Computer Virology., 2006.

ДОДАТКИ