

Міністерством освіти та науки України  
Прикарпатський національний університет імені Василя Стефаника  
Факультет математики та інформатики  
Кафедра комп'ютерних наук та інформаційних технологій

## **ДИПЛОМНА РОБОТА**

на здобуття другого (магістерського) рівня вищої освіти

на тему: «Використання нейронних мереж  
для розпізнавання касових документів»

Виконав: студент 2 курсу, групи КНМ-21

Спеціальності 122 Комп'ютерні науки

Бойко Б. Б.

Керівник: к.т.н. доцент Ровінський. В. А.

Рецензент

Івано-Франківськ – 2020 р.

### **Анотація**

Структура курсової роботи. складається зі вступу, трьох розділів, висновків, списку використаних джерел зі 11 найменувань, а також містить 7 рисунків. Загальний обсяг роботи складає 35 сторінок.

У першому розділі проводиться аналіз розвитку інформаційних технологій розпізнавання символів. В другому розділі проводиться розробка інформаційної технології розпізнавання символів. В третьому відбувається реалізація алгоритму розпізнавання символів.

Ключові слова: розпізнавання символів, штучна нейронна мережа, перцептрон, прихована марковська модель.

## Зміст

Вступ.....	4
1. Аналіз методів та технологій розпізнавання символів .....	6
1.1 Аналіз існуючих методів розв’язання задачі розпізнавання символів .....	6
1.2 Постановка задачі розпізнавання символів.....	10
1.3 Висновок.....	11
2. Розробка інформаційної технології розпізнавання символів .....	13
2.1 Розробка методу розв’язання задачі розпізнавання символів .....	13
2.2 Обґрунтування вибору виду нейронної мережі для розпізнавання символів	17
2.2.1 Нейронні мережі зі зворотним поширенням помилки .....	18
2.2.2 Нейронні мережі Хебба.....	20
2.2.3 Нейронні мережі Хопфілда .....	20
2.2.4 Нейронні мережі Хеммінга .....	21
2.2.6 Вибір нейронної мережі для інформаційної технології.....	23
2.3 Роль сегментації в обробці розпізнавання.....	24
2.4 Висновок.....	26
3. Реалізація розпізнавання символів в касових документах .....	27
3.1 Модель реалізації .....	27
3.2 Попередня обробка зображення.....	28
3.3 Виділення ознак .....	29
3.4 Навчання ПММ .....	30
3.5 Розпізнавання .....	31
Висновки.....	33
Список використаної літератури.....	34

## **Вступ**

### **Актуальність теми дослідження.**

З часом кількість інформації, яку споживає людина збільшилась в рази, це пов'язано з полегшенням доступу до інформації та її перенасичення в мережі та в житті загалом. Разом з цим виникає проблема запам'ятовуванням і зберіганням необхідної нам інформації.

Швидкість розвитку програмного забезпечення досить стрімка, але її впровадження може не встигати на запити сьогодення. В сучасному світі метою будь-якої програми – автоматизувати деякий вид діяльності людини, чи то великі розрахунки, чи виконання платежів.

Вид діяльності в якому програмні засоби все ж таки відстають від людини – це задачі розпізнавання та класифікації. Хоча вже і в цьому вдалось досягнути хорошого результату. Так задача розпізнавання тексту з графічного в символічне представлення спростивши для людини процес механічного вводу тексту зекономивши при цьому один з найцінніших для людини ресурсів – час.

Також немаловажливим є актуальне питання екології, адже перевезення касових документів має значний вплив на викиди транспортними засобами.

Отже, є необхідність в програмному забезпеченні, що міг би обробляти графічні текстові дані та з допомогою штучної нейронної мережі перетворював в текстовий матеріал. Таким чином у галузі розпізнавання символів актуальним є дослідження нових підходів класифікації для забезпечення коректної і швидкої її роботи

**Мета та завдання дослідження.** Метою дослідження магістерської кваліфікаційної роботи є збільшення точності розпізнавання символів в касових документах з використанням штучних нейронних мереж. Щоб досягнути поставленої мети необхідно виконати такі завдання:

- 1) проаналізувати процес розпізнавання символів;
- 2) розглянути існуючі методи вирішення задачі розпізнавання символів;
- 3) розробити структуру нейронної мережі;
- 4) сформулювати стадії інформаційної технології розпізнавання символів;
- 5) розробити алгоритм реалізації розпізнавання символів на основі нейронної мережі;
- 6) виконати реалізацію інформаційної технології розпізнавання символів на основі нейронної мережі;

**Об'єкт дослідження** – це процес розпізнавання друкованих символів.

**Предмет дослідження** – це методи розпізнавання символів на основі нейронних мереж.

**Методи дослідження:** системний аналіз та теорія нейронних мереж для реалізації інформаційної технології розпізнавання символів.

**Наукова новизна** одержаних результатів полягає в тому що запропонований метод дає змогу зменшити час навчання нейронної мережі для класифікації, зберігаючи достатньо високу точність.

**Практичне значення отриманих результатів.** Практична значимість отриманих результатів роботи полягає в можливості індексації і обліку касових документів, що значно спростить доступ до необхідної інформації.

# **1. Аналіз методів та технологій розпізнавання символів**

## **1.1 Аналіз існуючих методів розв'язання задачі розпізнавання символів**

Дослідники зацікавлені в розробці систем ОРС через численні потенційні можливості застосування в бізнесі та промисловості. Використання ОРС різниться залежно від областей застосування. Банківська справа - одна із широко відомих сфер, де ОРС використовується для автоматичної обробки чеків та інших форм. Напис на чеку можна миттєво відсканувати та розпізнати, щоб скоротити час очікування в банках. Системи ОРС дозволяють інструментам обробки форм отримувати та зчитувати відповідну інформацію з паперових форм. Медичним працівникам доводиться мати справу з великим обсягом форм, що містять важливу інформацію про пацієнтів. Корисно не відставати від усієї інформації, вкладаючи її в центральну базу даних цифровим способом, щоб мати змогу ефективно отримати доступ до інформації за необхідності. Для масштабних проєктів оцифрування потрібна ефективна система ОРС для перетворення мільйонів друкованих книг та документів у цифрові архіви. Цифрові архіви надають доступ для пошуку вміст, прості засоби резервного копіювання та усувають необхідність фізичного зберігання друкованих документів.

Системи ОРС перетворюють двовимірне зображення тексту, яке може містити машинний друкований або рукописний текст, в ідеалі будь-яким сценарієм, із його зображення в машиночитаний текст. Системи ОРС зазвичай працюють у конвеєрі, і до фактичного розпізнавання тексту існує декілька етапів [1]. Типова система ОРС може включати попередню обробку, аналіз макета, розпізнавання символів та моделювання мови. Попередня обробка, як правило, включає бінаризацію, видалення шуму, корекцію перекосів та додатково виявлення сценарію та орієнтації. Аналіз макета визначає текстові стовпці, текстові блоки, текстові рядки та порядок читання текстової сторінки. Розпізнавання символів відповідає за розпізнавання тексту, що міститься в текстовому рядку. Статистичне моделювання мови покращує результати

розпізнавання тексту шляхом інтеграції попередніх знань про мову, словниковий запас та сфера дії документа.

Методи розпізнавання символів використовують оптичні характеристики. Тобто існує декілька характерних для символу властивостей і система перевіряє наявність таких в зображенні. Такий пошук виконується досить швидко і має високі показники точності розпізнавання, однак має декілька суттєвих недоліків. По-перше для кожного символу і для кожного шрифту потрібно знайти та формалізувати десятки або і сотні ознак, при чому перевірити їх на наявність перетинів. По-друге виявлення завад або перешкод для розпізнавання лише виконується попередньою обробкою, а отже зникнення символу може призвести до втрати цієї інформації. Взагалі метод не придатний для розпізнавання рукописних символів, він може бути модифікований до поставленої задачі, однак буде мати дуже низькі показники якості розпізнавання.

Існує принципово інший підхід щодо задачі розпізнавання образів. Для розпізнавання можна використовувати кліткові автомати. Такі автомати – це набір локальних правил, що діють на дискретний світ. Правила діють в будь-якому місці дискретного світу однаково і використовують локальне оточення для зміни свого стану. Час в клітковому автоматі змінюється дискретними порціями змінюючи вигляд кліткового поля. Поле –  $N$ -вимірний простір розбитий на дискретні клітини за деякими регулярними правилами. Існує багато різновидів кліткових автоматів: стандартні кліткові автомати – основані на класичному визначенні даному вище; мобільні кліткові автомати – додатково містять одну активну клітину, яка змінює свій стан; тюрингові машини – активна клітина має декілька станів; системи підстановок – коли зміни проходять блоками клітинок і т. д.

Методики ґрунтуються на використанні бінарного зображення і обробці клітковими автоматами для виділення ознак конкретного символу. Такий метод вимагає ще більш складної підготовки, потрібно розробити систему правил для кліткових автоматів і до того ж потрібно виконати обробку одразу на всіх

автоматах, що вимагає дуже великого обсягу обчислень. Більш загальними є адаптивні методи розпізнання, основані на інтелектуальних методах класифікації. Одним з найвживаніших є використання нейронної мережі. Нейронна мережа являє собою систему з однотипних елементів обробки, які на основі вхідних даних формують вихідний сигнал. В залежності від параметрів елементів змінюється характер роботи мережі, цей факт використовується для використання системи для задач класифікації та машинного навчання. Існують методи побудови, методика визначення оптимальних параметрів мережі, таких як кількість шарів, розмір шарів, топологія, основні методи подання вхідного сигналу та їх переваги та недоліки.

Одним з найкращих (на практиці) методів подання вхідного сигналу є виділення знімків векторного поля зображення. На відміну від точкового знімку такому методу вистачає мережі меншого розміру і як наслідок меншого обсягу обчислень. Проте такий метод вимагає більш складної попередньої обробки. Нейронна мережа сама по собі виконує задачу класифікації. В багатовимірному просторі нейронна мережа виділяє області які відповідають деяким екземплярам класу.

Основними перевагами нейромережевих структур є їх можливість навчання. Більш того вони можуть досить вдало вирішувати класифікацію об'єкта який не очікувався при навчанні, оскільки апроксимують об'єкт у відповідність до відомих критеріїв.

Недоліками нейромережевих структур є низька точність розпізнавання та високі потреби на обрахунок (у випадку моделювання мережі на Фон Неймановській архітектурі, оскільки взагалі потреба в швидкості роботи процесора низька, однак важлива кількість процесорів). Також великим недоліком є низька точність.

Штучні нейронні мережі (ШНМ) спочатку були розроблені як математичний інструмент для моделювання можливостей обробки інформації людського мозку [2]. Хоча складна структура та механізм роботи людського



мозку не зрозумілі до кінця, можна моделювати механізм роботи вищого рівня біологічних нейронів у формі штучних нейронів. Основна структура ШНМ складається з невеликих процесорних блоків або вузлів, які з'єднані між собою зваженими з'єднаннями. У біологічному плані вузли представляють нейрони, а ваги з'єднань - силу синапсів між нейронами. Штучну нейронну мережу можна активувати, надаючи вхід для вхідних вузлів мережі. Вхідні вузли, як правило, не обробляють елементи і відповідають лише за отримання вхідних даних поза межами мережі. Вхід поширюється по всій мережі через зважені з'єднання. Ваги мережі були засвоєні під час контрольованого навчального процесу, і зважені зв'язки в основному представляють загальні знання системи розпізнавання. Активація мережі призвела до серії «сплесків» на вихідному рівні. Піки показують активації конкретних вихідних нейронів, які активуються на основі певних вхідних моделей.

За останні роки пропонується багато варіацій базової структури ШНМ [3]. Однією із широко використовуваних структур ШНМ є пряме поширення, в якій блоки розташовані шарами із з'єднаннями, що подаються вперед від одного шару до наступного. Добре відомі приклади нейронних мереж прямого поширення включають перцептрони, багат шарові перцептрони, радіально-базові функціональні мережі та згорткові нейронні мережі тощо. ЗНМ - це варіант багат шарових перцептронів, які надихнулись ранньою роботою Хубеля та Візеля над зоровою корою кота [4]. Вони виявили складне розташування місцевих чутливих клітин у зоровій системі кота. Клітини чутливі до невеликих субрегіонів вхідного простору, які називаються рецептивним полем, і нахилені до всього візуального поля. Зображення мають сильну двовимірну локальну структуру і пікселі, що просторово розташовані один до одного, мають високу кореляцію. Місцеві рецептивні поля найкраще підходять для використання місцевих кореляцій, присутніх на зображеннях. Перевагами просторової кореляції є вилучення та поєднання місцевих особливостей з точки зору ребер, точок чи кутів тощо. Перша реалізація згорткової нейронної мережі (ЗНМ), яка називається Неокогнітрон, була запропонована Фукусімою для розпізнавання рукописних

цифр. Неокогнітрон використовує ідею місцевих рецептивних полів (тобто кожен нейрон підключений лише до підрегіону, що відповідає деяким сусіднім нейронам у попередньому шарі) для вилучення місцевих особливостей із вхідних моделей.

Незважаючи на недоліки, нейронна мережа є одним з найпотужніших об'єктів для розпізнавання рукописних символів. Нейромережеві структури досить лояльно відносяться до незнайомих образів та намагаються класифікувати до найближчого знайомого представника. На практиці людина вводить досить схожий на оригінал символ, однак нелінійно додаючи деякі похибки. В разі розпізнавання друкованого тексту похибки або можна апроксимувати як лінійні або ж ні, якщо не зважати на оптичні завади та шуми. В обох випадках нейронна мережа здатна знайти деяку схожу апроксимацію, інші методи можуть зовсім не розпізнати символ.

Враховуючи описані вище методи та їх недоліки, можна зробити висновок, що задача розпізнавання все ще потребує більш швидкого та точного вирішення. Найбільш популярні та практичні програми використовують закономірності в символах для розпізнавання, що дає високі показники розпізнавання такі як швидкість або точність, однак можуть виявитись безпомічними при появі завади, або частковій втраті інформації.

## **1.2 Постановка задачі розпізнавання символів.**

Обчислювальні системи створювали з метою замінити людину в тих сферах діяльності, де людина виконує легку монотонну роботу. Сьогодні більшість звичних задач пов'язані з обчислення виконуються на комп'ютері. ПК здатен швидко виконувати прості арифметичні операції і з більшою точністю. Але деякі процеси вирішені не повністю. Особливо це стосується непростих задач, які потребують інтелектуальних методів вирішення. Одна з таких – розпізнавання тексту. Вже розроблено багато програмних додатків, які здатні розпізнавати тексти. Найкращі розробки використовують дуже прості методи, але завжди є вузькоспеціалізованими. Так існують додатки в яких для кожного типу шрифту розробляється своя система пошуку шаблонних ділянок. Такі методи дійсно

працюють, і працюють досить швидко і якісно. Але людина може досить легко перемогти такі системи по якості розпізнавання.

Людина використовуючи свій мозок та асоціативну пам'ять, розглядає зображення в цілому, використовує знання про контекст. Такі додаткові (хоча і незначні) затрати дозволяють досить легко справлятися з розпізнаванням слів, речень, навіть якщо частина зображення зіпсована.

Розпізнавання образів, або тексту використовується або може бути використано для оптимізації роботи в багатьох сферах. Наприклад для оцифрування літератури в бібліотеці, для сканування документів, для автоматичної перевірки. Хоча інформація для сканування може бути попередньо оброблена та підготовлена, існує безліч прикладів де використати попередню обробку досить складно або й взагалі неможливо.

Наприклад для автоматизації реєстрації пасажирів на літак використовуються автоматизовані системи, де для пошуку квитка виконується сканування паспорта. Така система має зручний інтерфейс та не ускладнює процедуру для пасажирів, проте дозволяє компанії використовувати меншу кількість персоналу. Такі системи вдало функціонують в багатьох авіакомпаніях.

Очевидно, що для отримання ефекту дуже важливо щоб якість роботи системи була максимально високою, оскільки неможливість розпізнати інформацію при зчитуванні, або можливі наслідки під час помилки можуть привести до отримання клієнтом невірною квитка. В інших випадках такі проблеми можуть стосуватись цінних паперів і мати ще потужніші наслідки.

### **1.3 Висновок**

У даному розділі була поставлена проблема недостатньої достовірності сучасного процесу розпізнавання символів. Були розглянуті основні методи та програмні засоби, які виконують розпізнавання символів та визначені їх наступні недоліки: 1) недостатня достовірність розпізнавання, яка полягає у малій відносній кількості правильно розпізнаних символів при наявності завад на

зображенні символів; 2) низька швидкість розпізнавання, яка викликана послідовним характером порівняння символу із базою еталонів. Для усунення цих недоліків було запропоновано використовувати метод розпізнавання символів, заснований на використанні штучних нейронних мереж.

## **2. Розробка інформаційної технології розпізнавання символів**

### **2.1 Розробка методу розв'язання задачі розпізнавання символів**

Розпізнавання символів являє собою дуже складну задачу в теоретичному та практичному сенсі, хоча багато живих організмів в тому числі і люди справляються з нею швидко і доволі точно. Досить важко створити штучну систему і технічно її реалізувати для того щоб ефективно виконувати даний процес. Під розпізнаванням образів розуміють співставлення властивостей об'єкта зображенню образу об'єкта.

Прикладом розпізнавання образів можуть являтися системи розпізнавання як тексту так і окремих символів, біометричних параметрів людини, штрих-кодів, номерів машин і т.д. Система розпізнавання тексту передбачає наявність на вході зображення з текстом (в заданому графічному форматі). На виході системи повинно бути сформовано текст. Загальна архітектура подібної системи складається з таких послідовних дій:

1. Поступаючи на вхід зображення повинно спочатку бути очищене від шумів та приведено до вигляду, що дозволяє ефективно виділяти символи та розпізнавати їх.
2. Система повинна розбити зображення на блоки тексту, основується на особливостях зображення.
3. Система повинна в загальному вигляді розбити блоки тексту на слова, а слова на символи, оскільки розпізнавання символів є більш простою задачею ніж розпізнавання цілого слова.
4. Після цього система виконує специфічні алгоритми для розпізнавання. Саме розпізнавання символу може виконуватися різними специфічними методами. Можна порівнювати з шаблоном, шукати блоки очікуваної для конкретного символу інформації використовувати нейромережу і т.д..

Після розпізнавання отримують множину символів з якої формується текст. Існує декілька основних систем розпізнавання тексту. Принцип роботи подібних систем базується на декількох стратегіях, але взагалі алгоритм в загальному вигляді базується на висуванні гіпотез з подальшою перевіркою їх. Причому від порядку висування гіпотез залежить швидкість алгоритму, тому в загальному випадку важливо знайти оптимальний порядок висування гіпотез, оснований на даних отриманих із зображення.

Оптичне розпізнавання символів зазвичай працює з растровим зображенням сторінки. При цьому більшість систем має шаблони створені для різних шрифтів та начерків. Програма розпізнає шрифт та виконує розпізнавання на основі шаблону для даного шрифту. В деяких випадках програмне забезпечення використовує числові значення частин символів щоб визначити новий шрифт.

Найпопулярніші програми розпізнавання тексту основані на пошуку специфічних особливостей символів. Такі системи складаються з сотень алгоритмів кожен з яких шукає ознаки ідентифікації конкретного символу. Такі системи інколи використовують принципи роботи нечітких множин для визначення ймовірності розпізнавання символу. Однак така система має ряд недоліків. Відсутність одного пікселя, або наявність регулярного чи нерегулярного шуму може дуже сильно зіпсувати якість розпізнавання.

Нейронна мережа може бути класифікатором в системі розпізнавання тексту. Цей класифікатор можна навчати, налаштовуючи коефіцієнти на елементах мережі таким чином, щоб в залежності від входу на виході з'являлись бажані результати. Нейронні мережі з успіхом можуть бути використані в системі розпізнавання образів. Суть дії мережі полягає в тому, що вона виділяє області в багатовимірному просторі всіх можливих вхідних даних. Кожна область, як приклад, може відповідати символу. Якщо після навчання подати на вхід мережі, точку з цього простору (чим і є будь-який вхідний сигнал), мережа спробує виявити до якої області належить сигнал. Як результат – мережа здатна видавати відповіді навіть на вхідні сигнали які досі мережі не зустрічались, при чому

результатом роботи буде область яка найближче до даної точки. Тому точність роботи мережі насправді залежить від якості навчання.

Один з недоліків нейронної мережі полягає в її розмірах. Зазвичай необхідні досить великі структури мережі. Проте така складність обумовлена традиційними методами обробки інформації. Сьогодні більшість комп'ютерів складаються з кількох дуже потужних процесорів, які дуже швидко виконують атомарні арифметичні дії. Нейронна мережа по суті складається з великої кількості простих елементів, які просто виконують роль суматора з пороговою функцією. Однак таких елементів дуже багато. Якщо порівнювати біологічну нейронну мережу людини з сучасним процесором, то математичні прості операції, наприклад, множення чи ділення чисел, людина виконує дуже повільно, хоча процесор виконує їх за мільярдні долі секунди. З іншого боку людині досить десятки секунди щоб зрозуміти що картина написала в стилі кубізму, що для двійкового процесора майже неможлива задача.

Простіший варіант реалізації нейронної мережі – моделювання такої на звичайному процесорі. Однак процесор обробляє інформацію послідовно, тому, хоча кожен елемент мережі незалежний і всі обрахунки можуть виконуватись паралельно, для опрацювання одного вхідного сигналу потрібен час пропорційний розміру мережі. Існує багато різновидів нейронних мереж, всі вони діляться на одношарові та багатшарові. Хоча один нейрон і здатний виконувати найпростіші процедури розпізнавання, сила нейронних обчислень виникає від сполук нейронів у мережах. Найпростіша мережа складається з групи нейронів, що утворюють шар, як показано в правій частині рис. 2.1. Вершини кола ліворуч служать лише для розподілу вхідних сигналів. Вони не виконують обчислень, і тому не будуть вважатися шаром. З цієї причини вони позначені малими колами, щоб відрізнити їх від обчислючих нейронів, позначених великими колами. Кожен елемент із множини входів  $X$  окремою вагою з'єднаний з кожним штучним нейроном, а кожен нейрон видає зважену суму входів в мережу.

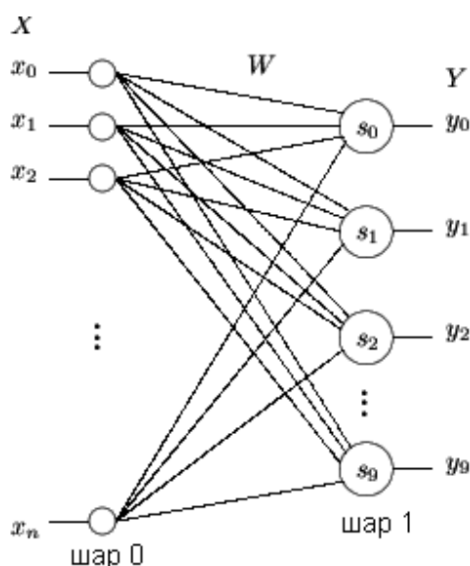


Рисунок 2.1 – Одношарова нейронна мережа

Великі й складні нейронні мережі як правило володіють великими обчислювальними можливостями. Хоча створено мережі всіх конфігурацій, які тільки можна собі уявити, пошарова організація нейронів копіює шаруваті структури певних відділів мозку. Такі багатошарові мережі володіють більшими можливостями, ніж одношарові.

Багатошарові мережі можуть утворюватися каскадами шарів. Вихід одного шару є входом для подальшого шару. Подібна мережа показана на рис. 2.2. Багатошарові мережі не можуть привести до збільшення обчислювальної потужності в порівнянні з одношаровою мережею лише в тому випадку, якщо активаційна функція між шарами буде лінійною. Обчислення виходу шару полягає в множенні вхідного вектора на першу вагову матрицю з подальшим множенням (якщо відсутня нелінійна активаційна функція) результуючого вектора на другу вагову матрицю  $-(XW_1)W_2$ . Так як множення матриць асоціативне, то  $X(W_1W_2)$ . Це показує, що двошарова лінійна мережа еквівалентна одному шару з ваговою матрицею, рівною добутку двох вагових матриць. Отже, будь-яка багатошарова лінійна мережа може бути замінена на еквівалентну одношарову мережу, проте, одношарові мережі досить обмежені за своїми обчислювальними можливостями. Таким чином, для розширення можливостей



мереж в порівнянні з одношаровою мережею необхідна нелінійна активаційна функція.

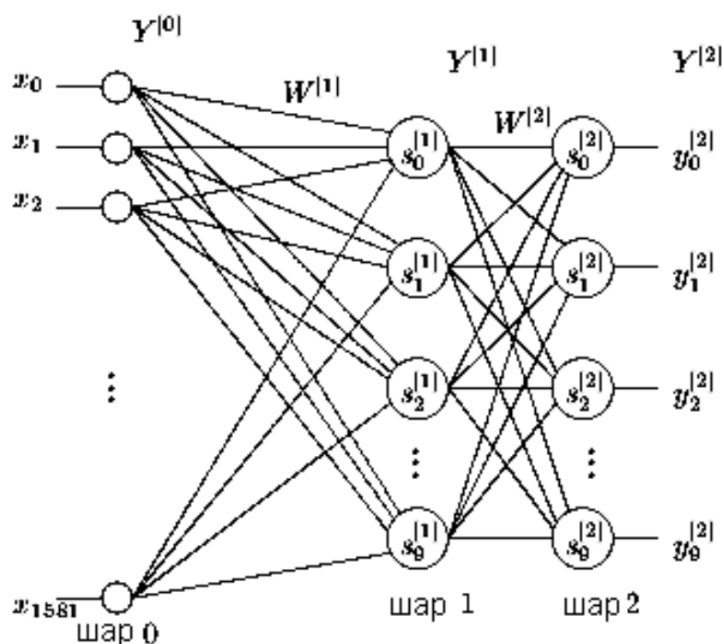


Рисунок 2.2 – Багатошарова нейронна мережа

## 2.2 Обґрунтування вибору виду нейронної мережі для розпізнавання символів

Нейронна мережа дійсно найкращий об'єкт для реалізації адаптивного рішення задачі класифікації. Однак, для ефективного розв'язання поставленої задачі потрібно обрати деякий вид мережі, проаналізувати існуючі варіанти реалізації нейромережі. Було вибрано 5 популярних варіанти реалізації нейромережі: нейронна мережа зі зворотним поширенням помилки, нейронна мережа Хебба, нейронна мережа Хопфілда, нейронна мережа Хеммінга та згорткова нейронна мережа. Аналіз декількох варіантів дозволить більш детально розглянути математичні засади, які лежать в процесах навчання нейромережі, і прийняти правильне рішення щодо використання конкретної реалізації.

## 2.2.1 Нейронні мережі зі зворотним поширенням помилки

Коли в мережі тільки один шар, алгоритм її навчання з учителем очевидний, так як правильні вихідні стани нейронів єдиного шару завідомо відомі, і підстроювання синаптичних зв'язків йде в напрямку, що мінімізує помилку на виході мережі. За цим принципом будується, наприклад, алгоритм навчання одношарового перцептрона. У багатошарових ж мережах оптимальні вихідні значення нейронів усіх шарів, крім останнього, як правило, не відомі, і перцептрон з двома і більше шарами вже неможливо навчити, керуючись тільки величинами помилок на виходах ШНМ. В мереж, розглянутих вище, не було зворотних зв'язків, тобто сполук, що йдуть від виходів деякого шару до входів цього ж шару або попередніх шарів. Цей спеціальний клас мереж, які називаються мережами без зворотних зв'язків або мережами прямого розповсюдження. В мережах без зворотних зв'язків немає пам'яті, їх вихід повністю визначається поточними входами і значеннями ваг. Один з варіантів вирішення цієї проблеми - розробка наборів вихідних сигналів, відповідних входним, для кожного шару ШНМ, що, звичайно, є дуже трудомісткою операцією і не завжди здійсненою. Другий варіант - динамічне підстроювання вагових коефіцієнтів синапсів, в ході якого вибираються, як правило, найбільш слабкі зв'язки і змінюються на малу величину в ту чи іншу сторону, а зберігаються тільки ті зміни, які спричинили зменшення помилки на виході всієї мережі. Третій варіант - поширення сигналів помилки від виходів ШНМ до її входів, у напрямку, зворотньому прямим поширенню сигналів у звичайному режимі роботи.

У процесі навчання може виникнути ситуація, коли великі позитивні або негативні значення вагових коефіцієнтів змістять робочу точку на сігмоїдах багатьох нейронів в область насичення. По-друге, застосування методу градієнтного спуску не гарантує, що буде знайдено глобальний, а не локальний мінімум цільової функції. Ця проблема пов'язана ще з однією, а саме - з вибором величини швидкості навчання. Доказ збіжності навчання в процесі зворотного поширення заснований на похідних, тобто збільшенні ваг, отже, швидкість

навчання повинна бути нескінченно малою, проте в цьому випадку навчання відбуватиметься неприйнятно повільно. З іншого боку, занадто великі корекції ваг можуть призвести до постійної нестійкості процесу навчання. Тому як  $\eta$  зазвичай вибирається число менше 1, але не дуже маленьке, наприклад, 0.1, і воно може поступово зменшуватися в процесі навчання. Крім того, для виключення випадкових попадань в локальні мінімуми іноді, після того як значення вагових коефіцієнтів стабілізуються,  $\eta$  короткочасно сильно збільшують, щоб почати градієнтний спуск з нової точки. Якщо повторення цієї процедури кілька разів приведе алгоритм в один і той же стан ШНМ, можна більш-менш впевнено сказати, що знайдений глобальний максимум, а не якийсь інший.

Є й інший метод виключення локальних мінімумів, а заодно і паралічу ШНМ, що полягає в застосуванні стохастичних ШНМ.

Алгоритм навчання нейронної мережі за допомогою процедури зворотного поширення має на увазі наявність якоїсь зовнішньої ланки, що надає мережі крім вхідних також і цільові вихідні образи. Алгоритми, що користуються подібної концепцією, називаються алгоритмами навчання з учителем. Для їх успішного функціонування необхідна наявність експертів, що створюють на попередньому етапі для кожного вхідного образу еталонний вихідний. Наприклад, навчання людського мозку, на перший погляд, відбувається без учителя: на зорові, слухові, тактильні та інші рецептори надходить інформація ззовні, і всередині нервової системи відбувається якась самоорганізація. Проте, не можна заперечувати і того, що в житті людини не мало вчителів - і в буквальному, і в переносному значенні, які координують зовнішні впливи.

Головна риса, що робить навчання без учителя привабливим, - це його "Самостійність". Процес навчання, як і у випадку навчання з учителем, полягає в підстроювання ваг синапсів. Деякі алгоритми змінюють структуру мережі, тобто взаємозв'язку нейронів або навіть їх кількість, такі перетворення називаються - самоорганізацією. Очевидно, що підстроювання синапсів може проводитися тільки на підставі інформації, доступної в нейроні, тобто його стану та вже

існуючих вагових коефіцієнтів. Виходячи з цього міркування і за аналогією з відомими принципами самоорганізації нервових клітин, побудовані алгоритми навчання Хебба.

### **2.2.2 Нейронні мережі Хебба**

Згідно з поглядами Хебба нервові клітини мозку з'єднані одне з одним великою кількістю прямих та зворотних збуджуючих зв'язків і утворюють нейронну мережу. Кожен нейрон здійснює просторово-часову сумачію сигналів, що приходять до нього, від порушених нейронів, визначаючи потенціал на своїй мембрані. Коли потенціал на мембрані перевищує порогове значення, нейрон збуджується. Нейрон має рефрактерність і втому. Ефективність зв'язків може змінюватися в процесі функціонування мережі, підвищуючись між одночасно порушеними нейронами. Це призводить до об'єднання нейронів у клітинні ансамблі - групи клітин, які частіше всього порушувалися разом, і до відокремлення ансамблів один від одного. При порушенні достатньої частини ансамблю він збуджується цілком. Різні ансамблі можуть перетинатися: один і той же нейрон може входити в різні ансамблі. Електрична активність мозку обумовлена послідовним збудженням окремих ансамблів.

### **2.2.3 Нейронні мережі Хопфілда**

Серед різних конфігурацій штучних нейронних мереж зустрічаються такі, при класифікації яких за принципом навчання, строго кажучи, не підходять ні навчання з учителем, ні навчання без учителя. У таких мережах вагові коефіцієнти синапсів розраховуються тільки один раз перед початком функціонування мережі на основі інформації про оброблювані дані, і все навчання мережі зводиться саме до цього розрахунку. З одного боку, пред'явлення апріорної інформації можна розцінювати, як допомога вчителя, але з іншого - мережа фактично просто запам'ятовує зразки до того, як на її вхід надходять реальні дані, і не може змінювати свою поведінку, тому говорити про ланку зворотного зв'язку зі "світом" (вчителем) не доводиться. З мереж з подібною

логікою роботи найбільш відомі мережа Хопфілда та мережа Хеммінга, які зазвичай використовуються для організації асоціативної пам'яті. Структурна схема мережі Хопфілда наведена на рис. 2.3. Вона складається з єдиного шару нейронів, число яких є одночасно числом входів і виходів мережі. Кожен нейрон зв'язаний синапсами з усіма іншими нейронами, а також має один вхідний синапс, через який здійснюється введення сигналу. Вихідні сигнали зазвичай утворюються на аксонах. Завдання, яке вирішується цією мережею в якості асоціативної пам'яті, як правило, формулюється таким чином. Відомий деякий набір двійкових сигналів (Зображень, інших даних, що описують якісь об'єкти або характеристики процесів), які вважаються зразковими. Мережа повинна вміти з довільного неідеального сигналу, поданого на її вхід, виділити ("згадати" по частковій інформації) відповідний зразок (якщо такий є) або "дати висновок" про те, що вхідні дані не відповідають жодному зі зразків.

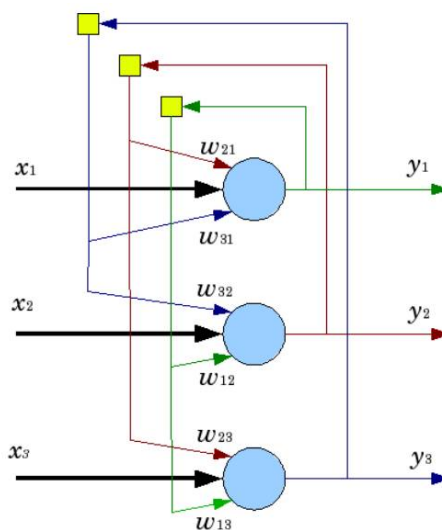


Рисунок 2.3 – Структурна схема мережі Хопфілда

## 2.2.4 Нейронні мережі Хеммінга

Коли немає необхідності, щоб мережа в явному вигляді видавала зразок, тобто достатньо отримувати номер зразка, асоціативну пам'ять успішно реалізує мережа Хеммінга. Дана мережа характеризується, в порівнянні з мережею

Хопфілда, меншими витратами на пам'ять і обсягом обчислень, що стає очевидним з її структури. Мережа Хеммінга (рис. 2.4) складається з двох шарів. Перший і другий шари мають по  $m$  нейронів, де  $m$  - число зразків. Нейрони першого шару мають по  $n$  синапсів, з'єднаних з входами мережі (утворюючими фіктивний нульовий шар). Нейрони другого шару пов'язані між собою інгібіторними (негативними зворотними) синаптичними зв'язками. Єдиний синапс з позитивним зворотним зв'язком для кожного нейрона з'єднаний з його ж аксоном.

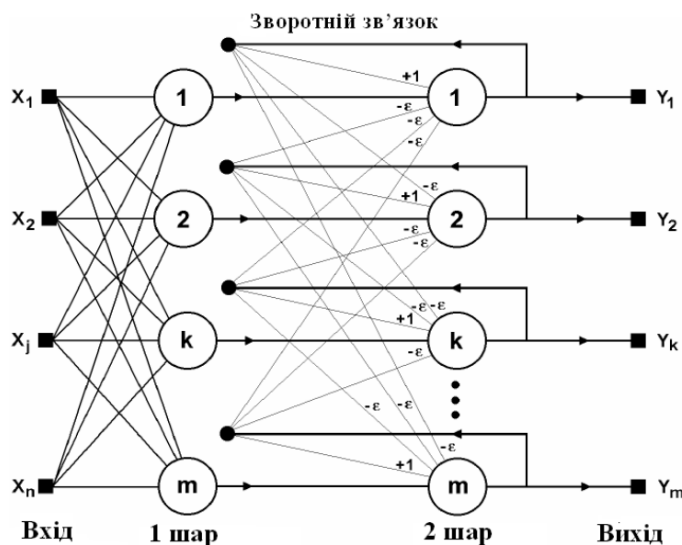


Рисунок 2.4 – Структурна схема мережі Хеммінга

Ідея роботи мережі полягає в знаходженні відстані Хеммінга від способу що тестується, до всіх зразків. Відстанню Хеммінга називається число бітів що відрізняються в двох бінарних векторах. Мережа має вибрати зразок з мінімальною відстанню Хеммінга до невідомого вхідного сигналу, в результаті чого буде активізований тільки один вихід мережі, відповідний цьому зразку.

### 2.2.5 Згорткова нейронна мережа

Згорткові нейронні мережі складаються з безлічі шарів штучних нейронів. Штучні нейрони, є математичними функціями, які обчислюють зважену суму кількох входів і виводять значення активації. Поведінка кожного нейрона

визначається його вагою. При подачі значень пікселів штучні нейрони CNN виділяють різні візуальні особливості.

Коли ви вводите зображення в ConvNet, кожен з його шарів генерує кілька карт активації. Карти активації виділяють відповідні особливості зображення. Кожен з нейронів бере в якості вхідного сигналу ділянку пікселів, множить їх значення кольору на його вагу, підсумовує їх і запускає через функцію активації.

Перший шар CNN зазвичай виявляє основні ознаки, такі як горизонтальні, вертикальні та діагональні ребра. Вихід першого шару подається як вхід наступного шару, який витягує більш складні функції, такі як кути та комбінації ребер. По мірі проникнення в згорткову нейронну мережу шари починають виявляти функції вищого рівня, такі як об'єкти, обличчя тощо.

Операція множення значень пікселів на ваги та їх підсумовування називається «згорткою» (звідси і назва згорткової нейронної мережі). CNN зазвичай складається з декількох згорткових шарів, але він також містить інші компоненти. Кінцевий рівень CNN - це класифікаційний рівень, який приймає вихідні дані остаточного рівня згортки як вхідні дані (пам'ятайте, що вищі рівні згортки виявляють складні об'єкти).

На основі карти активації остаточного рівня згортки рівень класифікації видає набір оцінок достовірності (значення від 0 до 1), які визначають, наскільки ймовірно, що зображення належить до “класу”.

## **2.2.6 Вибір нейронної мережі для інформаційної технології**

Розглянувши п'ять типи мереж, оберемо найоптимальнішу для реалізації в системі розпізнавання образів. Для цього розглянемо найважливіші для нас критерії.

Оскільки, система повинна працювати максимально ефективно, потрібна мережа мінімального розміру, оскільки навчання мережі є процесом складним з точки зору обчислень.

Перший критерій швидкодія. Навчання мережі повинно вимагати якомога менше операцій. Для збільшення швидкодії можна було б додати додаткову вимогу: можливість проведення навчання на цілочисельній арифметиці. Однак такий підхід зменшить точність рішення, що негативно впливає на другий критерій: якість розпізнавання. Отже, зосередимось лише на ефективності з точки зору мінімальності операцій ділення, множення та додавання (віднімання, оскільки з точки зору комп'ютера ці дві операції еквівалентні по швидкодії).

Для першого критерію найкращим типом мережі буде звичайна мережа зі зворотнім розповсюдженням помилки. Це обґрунтовується тим, що навчання можна оптимізувати обравши функцію активацію з нескладною похідною. Таким чином ми зменшуємо кількість арифметичних операцій процесора. Інші типи мережі мають занадто складну структуру, що тягне за собою додаткові витрати процесорного часу на розгляд топології. Тому в якості методу навчання обрано алгоритм навчання з вчителем, а в якості типу мережі – згорткову нейронну мережу.

Однак використання лише одношарової мережі веде до зменшення якості розпізнавання і навчання. Так як розпізнавання касових документів розуміє під собою не надто варіативні зміни символів адже більшість з них використовують або одні і ті ж, або схожі шрифти, тому, для реалізації обрано згорткову нейронну мережу з алгоритмом навчання: навчання з вчителем.

### **2.3 Роль сегментації в обробці розпізнавання**

Сегментація символів - це операція, яка має на меті розкласти зображення послідовності символів на підобрази окремих символів. Це один із процесів прийняття рішень в системі оптичного розпізнавання символів (англ. optical character recognition - OCR). Його рішення про те, що візерунок, виділений із зображення, є символом (або якоюсь іншою ідентифікованою одиницею), може бути правильним чи неправильним. Неправильно досить часто робити великий внесок у рівень помилок системи. У тому, що можна назвати "класичним"



підходом до OCR, сегментація є початковим етапом у триступеневій процедурі: Дано початкову точку в зображенні документа:

1. Знаходження наступного зображення символу.
2. Отримання особливих атрибутів зображення символу.
3. Знаходження відповістей даного набору символів, атрибути яких найкраще відповідають вхідним, і вивести його збіг.

Ця послідовність повторюється, поки не буде знайдено додаткових зображень символів. Реалізація кроку 1, етапу сегментації, вимагає відповіді на просто поставлене запитання: "Що являє собою символ?". Символ - це шаблон, що нагадує один із символів, розроблених системою. Але для визначення такої подібності візерунок повинен бути сегментований із зображення документа. Кожен етап залежить від іншого, і в складних випадках парадоксальним є пошук шаблону, який би відповідав члену алфавіту розпізнавання системи символів, не включаючи в процес детальних знань про структуру цих символів. Крім того, рішення щодо сегментації не є місцевим рішенням, незалежним від попередніх та наступних рішень. Створення гарного збігу з символом бібліотеки необхідне, але недостатнє для надійного розпізнавання. Тобто, поганий збіг на пізнішому зразку може поставити під сумнів правильність поточного результату сегментації / розпізнавання. Навіть серію задовільних збігів шаблонів можна вважати неправильною, якщо контекстні вимоги до вихідних даних системи не задовольняються. Наприклад, послідовність букв "cl" часто може дуже нагадувати "d", але, як правило, такий вибір не буде представляти собою контекстуально підходящий результат. Таким чином видно, що рішення щодо сегментації взаємозалежно від місцевих рішень щодо схожості форм та від глобальних рішень щодо прийнятності в контексті. Це речення підсумовує вдосконалення процесів сегментації символів за останні 40 років або близько того. Спочатку дизайнери прагнули провести сегментацію відповідно до "класичної" послідовності, перерахованої вище. Як швидший, потужніший електронний схемотехніка заохочувала застосовувати OCR до більш складних документів, дизайнери

зрозуміли, що крок 1 не можна відокремити від інших аспектів процесу визнання. Насправді дослідники вже багато років усвідомлюють обмеження класичного підходу. Дослідники в 1960-х і 1970-х роках відзначали, що сегментація спричиняє більше помилок, ніж викривлення форми при читанні необмежених символів, будь то ручний або машинний друк. Проблема часто маскувалася в експериментальних роботах шляхом використання баз даних з добре сегментованими шаблонами або шляхом сканування рядків символів, надрукованих із додатковим інтервалом. У комерційних заявках були встановлені жорсткі вимоги до підготовки документів. На початку 1980-х років працівники почали заохочувати відновлення дослідницького інтересу, щоб дозволити розповсюдження OCR на менш обмежені документи.

## **2.4 Висновок**

У даному розділі було виконано теоретичне обґрунтування розробки інформаційної технології розпізнавання символів. Було обґрунтовано вибір штучних нейронних мереж як теоретичної основи для реалізації інформаційної технології, нейронної мережі, що забезпечує найвищу достовірність та швидкість розпізнавання символів та роль сегментації в розпізнаванні символів.

### **3. Реалізація розпізнавання символів в касових документах**

#### **3.1 Модель реалізації**

Приховані моделі Маркова (ПММ) довели свою успішність у багатьох програмах машинного навчання та розпізнавання образів для аналізу та моделювання послідовних даних [6]. Дослідники розробили ефективні системи розпізнавання образів для розпізнавання мови за допомогою ПММ [7]. На даний час ПММ вважаються найсучаснішими в статистичних методах для забезпечення автоматичного розпізнавання мови.

Успішне застосування ПММ для дослідження розпізнавання мови забезпечило багато корисних аспектів технології для створення надійної системи OCR. Наприклад, оптична система розпізнавання символів може бути реалізована навіть без надання інформації про сегментацію символів. Незалежність мови - ще один атрибут, і та сама система може бути застосована до інших мов без будь-яких або незначних модифікацій. Сьогодні ПММ вважаються потужним інструментом у дослідженнях розпізнавання тексту, і багато корисних систем були розкриті раніше. Ще однією можливою перевагою ПММ, є розпізнавання тексту з низькою роздільною здатністю..

Зазвичай системи OCR розробляються для певного сценарію чи мови, і вони можуть розпізнавати символи, що належать до цього конкретного сценарію чи мови. Сценарій може бути визначений як набір символів, що використовується для графічного представлення певної мови або групи мов. Мови у світі набираються в багатьох різних сценаріях. Сценарій може використовуватися лише однією мовою, або ним можуть користуватися декілька мов. Зазвичай, у багатомовному середовищі, системи OCR вимагають розпізнавання символів незалежно від класу сценарію. Однак побудова системи OCR, яка зможе читати символи з усіх сценаріїв, є дуже складною. Грубим рішенням було б навчити класифікатор OCR на декількох сценаріях, додавши окремі символи з усіх сценаріїв у процесі навчання. Однак це призвело б до більшої кількості помилок

класифікації через збільшення кількості класів символів. Крім того, функції, необхідні для розпізнавання символів, зазвичай залежать від структурних властивостей написання, яке, як правило, відрізняється від одного сценарію до іншого. Іншим рішенням є поєднання класифікаторів рівня символів або слів для різних мов або сценаріїв, а розпізнавання певного символу чи слова здійснюється за допомогою відповідного класифікатора. Але для цього потрібне попереднє знання сценарію для застосування відповідних класифікаторів.

### **3.2 Попередня обробка зображення**

Попередня обробка - це майже один з основних кроків у кожній системі розпізнавання тексту. Зазвичай він використовується для усунення шуму та різних варіацій даних. Це може включати бінаризацію, видалення шуму, корекцію перекосів, покращення зображення, нормалізацію даних тощо. Попередня обробка в основному складається з бінаризації та видалення шуму. Бінаризація. Повна експериментальна установка, яка використовується в цьому методі, базується на бінаризованих зображеннях. Різні сучасні методи бінаризації можна класифікувати на дві групи: глобальна бінаризація та локальна. Глобальна бінаризація визначає єдиний поріг для цілого зображення, тоді як локальна бінаризація обчислює поріг для кожного окремого пікселя на основі інформації про сусідній. Загалом, локальна бінаризація працює краще, ніж глобальна бінаризація, при різних типах деградації зображень документа, таких як нерівномірне затінення або розмиття тощо. Однак, локальні методи бінаризації повільніші, ніж глобальні методи бінаризації. Бінаризовані зображення можуть містити невеликий шум (сіть і перець) або великий шум (об'єднання меж символів). Шум усувається за допомогою евристики.



Рисунок 3.1 – Результат процесу бінаризації

### 3.3 Виділення ознак

ПММ працює над послідовними даними та зображеннями текстових рядків можуть бути перетворені у послідовне подання, переміщуючи розсувне вікно вздовж зображень текстових рядків у напрямку написання. Під час цього процесу із зображення текстового рядка витягуються невеликі вертикальні фрагменти або кадри. Рамки можуть певною мірою перекриватися залежно від руху вікна вздовж текстового рядка. Послідовність кадрів зображень забезпечує основу для подальшого вилучення об'єктів. Поточний метод також використовує розсувне вікно для оцінки двох різних характеристик піксельної шкали сірого та функції інтенсивності на основі градієнта - для розпізнавання зображених на екрані зображених текстових рядків.

Функції пікселів у шкалі сірого - це вихідні значення пікселів, отримані із введених зображень рядків тексту. Вхідні рядки тексту спочатку нормуються. Вікно шириною в один піксель переміщується по нормалізованому рядку тексту, а вектор ознак складається шляхом підбору значень пікселів кожного вікна. Метод обчислює 20 функцій на вікно, оскільки вхідні зображення нормуються до висоти

лінії 20 пікселів. Ці функції на основі пікселів збираються для кожного вхідного зображення рядкового тексту та пізніше використовуються для навчання ПММ разом із їх транскрипціями ASCII.

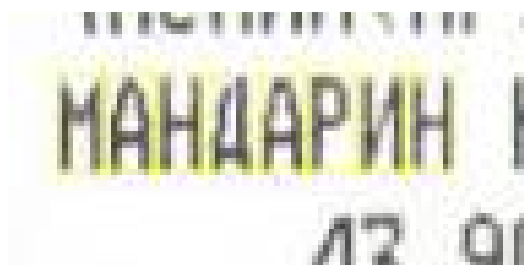


Рисунок 3.2 – Поділ на окремі символи

### 3.4 Навчання ПММ

Запропонована система ОРС моделює кожен символ із прихованою марковською моделлю безперервної щільності з різними станами. Переходи з одного стану в інший здійснюються зліва направо. Кожен стан у моделі асоціюється з розподілом вихідних ймовірностей за ознаками. Розподіли ймовірностей на виході моделюються як зважена сума Гауса, яку також називають гауссовими сумішами. У цьому методі гауссова суміш параметризується за допомогою засобів і дисперсій компонента Гауса і ваги кожного гауссова суміші. Кількість станів і допустимі переходи з одного стану в інший є параметрами системи. Значення цих параметрів можна регулювати на початку тренувального процесу. Кількість станів, які є адекватними для кожної моделі символів, залежить від горизонтальної мінливості кожного символу. Кількість станів може варіюватися в залежності від однієї моделі символів, однак для простоти кожен символ моделюється з фіксованою кількістю станів. У роботі використовуються 4 стани, зліва направо, ПММ з топологією. Кількість гауссів у стані визначається емпіричним шляхом. Спочатку обираються 2 компоненти Гауссової суміші на стан, і значення подвоюється після кожної ітерації тренування. Хороший показник розпізнавання отримується з 256 гауссів за стан. Алгоритм ітеративно вирівнює вектори об'єктів із моделями символів, щоб отримати

максимальні оцінки ймовірності параметрів ПММ. Навчання виконується з використанням навчального набору текстових рядків та відповідних основних файлів істинності.

### 3.5 Розпізнавання

Під час фази розпізнавання рядки вводу тексту попередньо обробляються, а функції витягуються, як описано у вищевказаних розділах. Процес розпізнавання шукає послідовність моделей символів, яка має найбільшу ймовірність генерування спостережуваної послідовності векторів ознак. Процес вимагає навчених моделей символів, статистичної мовної моделі та можливого лексикону слів чи символів. Вибір лексики та мовної моделі не обов'язковий, і їх використання, як правило, призводить до нижчого рівня помилок. Символьні ПММ комбінуються, щоб надати складні моделі для повних рядків тексту з використанням ергодичної топології ПММ. В ергодичній топології будь-яку модель можна отримати з будь-якої іншої моделі з кінцевою кількістю переходів. Розпізнавання тексту машинного друку вважається відносно простим завданням. Але це вірно лише в тому випадку, якщо основні зображення документів дуже чисті та відскановані. Але машинно роздрукований ОРС стає більш складним за наявності шуму та використання різних шрифтів, розмірів шрифтів та стилів шрифтів. Запропонований метод також оцінюється при розпізнаванні деградованих машинних друкованих зображень документів. Надрукований набір символів, використаний у цій роботі, містить 95 класів символів, включаючи алфавіти верхнього та нижнього регістру, цифри, розділові знаки та спеціальні символи. Помічено, що ПММ добре піддаються навчанням з різними варіантами часу навчання з рівномірними варіаціями, що відображається їх ефективністю при

розпізнаванні синтетичних текстових зображень, відтворених на екрані.

ТВОРОЖНА МАСА ГАВРЮША  
 КУРАГА 18% /400Г<ШТ>  
 25,20\* 1 = 25,20-А  
 ЧАСНИК<КГ> 54,90\* 0,270= 14,82-А  
 МАНДАРИН КЛЕМЕНТИН<КГ>  
 43,90\* 0,684 = 30,03-А  
**СУМА 70,05 ГРН**

творожна маса гаврюша  
 курага 18% /дог\_шт>  
 25,20\* 1 = 25,20\_0  
 часник<кг> 54,90\* 0,270% 14,82\_А  
 мандарин клементин<кг> 43,90\* 0,684 с 30,03\_А  
 (сума 70,05 грн)  
**СУМА 70,05 ГРН**

Рисунок 3.3 – Результат процесу розпізнавання



## **Висновки**

Було доведено, що розробка даної інформаційної технології та програми розпізнавання символів є актуальною, а також доцільною у порівнянні даної розробки з аналогами. Результати аналізу предметної області, в якому розглянуті сучасні методи розпізнавання образів, сучасні методи попередньої обробки зображень, методи подання зображень до нейромереж та основні топології нейронних мереж з алгоритмами їх навчання, дозволили обґрунтувати вибір нейронної мережі.

## Список використаної літератури

1. Thomas M. Breuel. The OCRopus open source OCR system. In Proc. of SPIE Document Recognition and Retrieval XV, volume 6815, 2008
2. Christopher M. Bishop. Neural Networks for Pattern Recognition. Oxford University Press, Inc., New York, NY, USA, 1995.
3. S. S. Haykin. Neural networks: a comprehensive foundation. NJ: Prentice Hall, Englewood Cliffs, New Jersey, USA, 2007
4. David H Hubel and Torsten N Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. The Journal of physiology, 160(1):106, 1962.
5. Горелик А.Л. Методы распознавания / А.Л. Горелик, В.А. Скрипник. – М. : Высшая школа, 1989. – 232 с
6. Thomas Plötz and Gernot A. Fink. Markov models for offline handwriting recognition: a survey. Int. Jour. Doc. Anal. Recognit., 12(4):269–298, November 2009
7. T.-H. Su, T.-W. Zhang, H.-J. Huang, and Y. Zhou. HMM-Based Recognizer with Segmentation-free Strategy for Unconstrained Chinese Handwritten Text. In Proceedings of the Ninth International Conference on Document Analysis and Recognition, pages 133–137, 2007.
8. Фукунага К. Введение в статистическую теорию распознавания / К.Фукунага. – М.: Наука, 1979. – 512 С.
9. Алексеев А. Алгоритм розпізнавання символів на основі структурного підходу / А. Алексеев, В. Заяць, Д. Іванов // Вісник Нац. ун-ту „Львівська політехніка” „Комп'ютерна інженерія та інформаційні технології”. – 2002. – № 468. – С.129 - 133.
10. Ф.Уосермена «Нейрокомп'ютерна техніка: Теорія і практика» Переклад українською, І.Ю. Юрчак, 2001.

11. Ю.И.Журавлев, В.В.Рязанов, О.В.Сенько Распознавание. Математические методы. Программная система. Практические применения. — М.: Фазис, 2005