

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ
Навчально-науковий інститут Комп'ютерних інформаційних
технологій
Кафедра Комп'ютеризованих систем управління

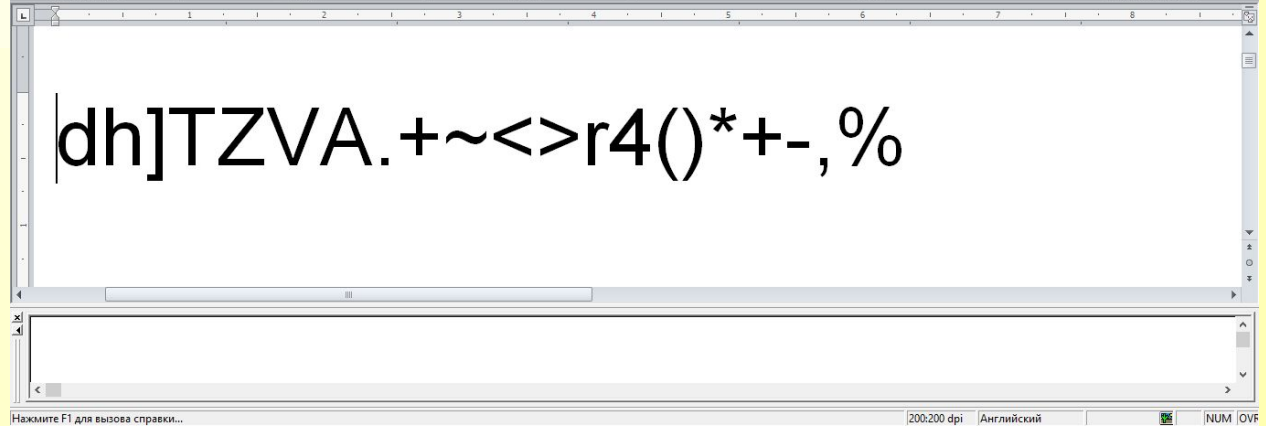
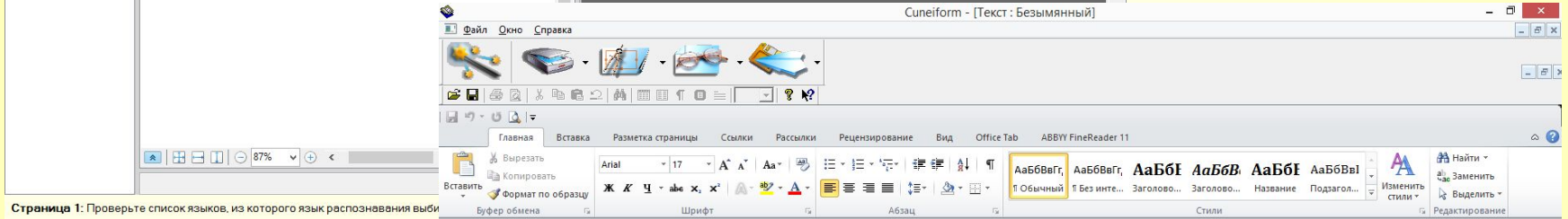
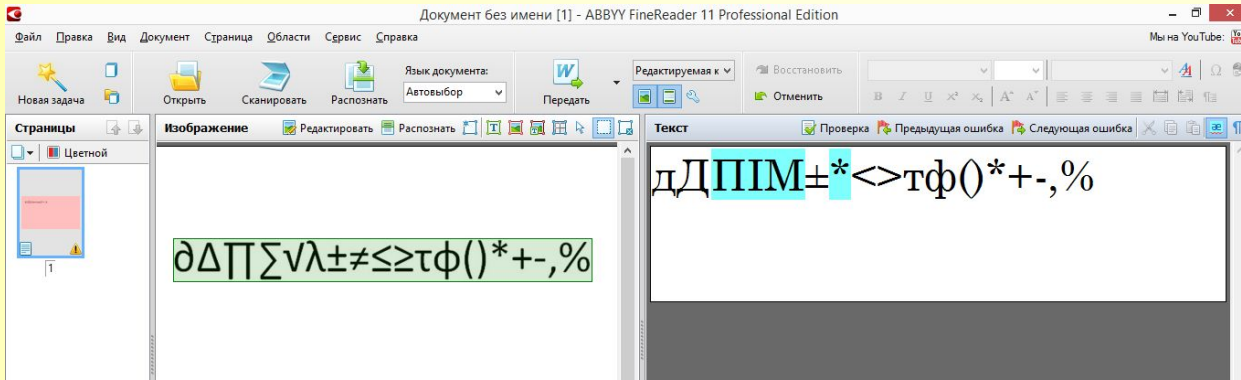
Дипломний проект
Тема: «Програмний модуль розпізнавання тексту з
растрових зображень»

Виконавець: Назаренко О.І.
Керівник: Нечипорук О.П.

Актуальність

Методи розпізнавання тексту	Переваги	Недоліки
Шаблонні методи	Достатньо висока точність розпізнавання дефектних символів (склеєних або розірваних)	Неможливість розпізнати шрифт, який хоч трохи відрізняється від закладеного в систему
Структурні, або топологічні, методи	Висока точність(≈95%) та швидкодія (1000 образів/с)	Ідентифікація знаків, що мають дефекти
Ознакові методи	Простота реалізації, хороша узагальнююча здатність, стійкість до зміни форми символів, низька кількість відмов, висока швидкодія	Нестійкість до дефектів зображення, на етапі вилучення ознак відбувається безповоротна втрата частини інформації про символ
Методи на основі нейронної мережі	Стійкість до шумів вхідних даних, адаптація до змін, відмовостійкість, висока швидкодія	Дають приблизну відповідь

Актуальність



- 1. Об'єкт проектування** – розпізнавання тексту з растрових зображень.
- 2. Предмет проектування** – програмний модуль розпізнавання тексту з растрових зображень.
- 3. Мета дипломного проекту** – розробити програмний модуль розпізнавання тексту з растрових зображень для розпізнавання спеціальних математичних символів та перенесення результатів розпізнавання до текстового редактору *Microsoft Office Word*.

Завдання проектування

1. Проаналізувати сучасні програмні рішення розпізнавання тексту з растрових зображень.
2. Проаналізувати методи розпізнавання тексту з растрових зображень.
3. Розробити структуру програмного модуля розпізнавання тексту з растрових зображень.
4. Реалізувати алгоритм розпізнавання спеціальних текстових символів на основі нейронної мережі.
5. Реалізувати алгоритм передачі спеціальних текстових символів в *MS Office Word*.
6. Проаналізувати роботу програмного модуля.

Структура програмного модуля

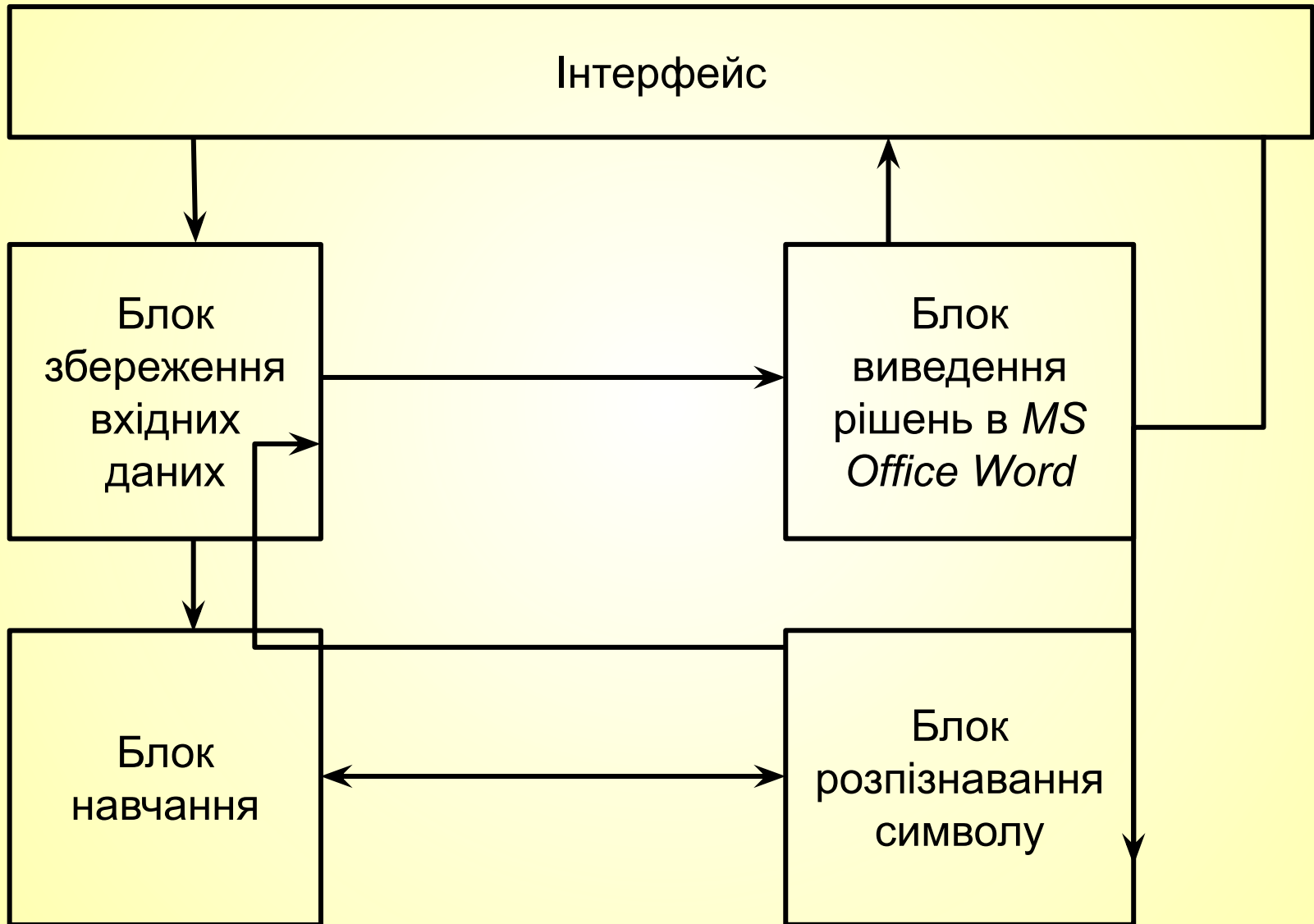
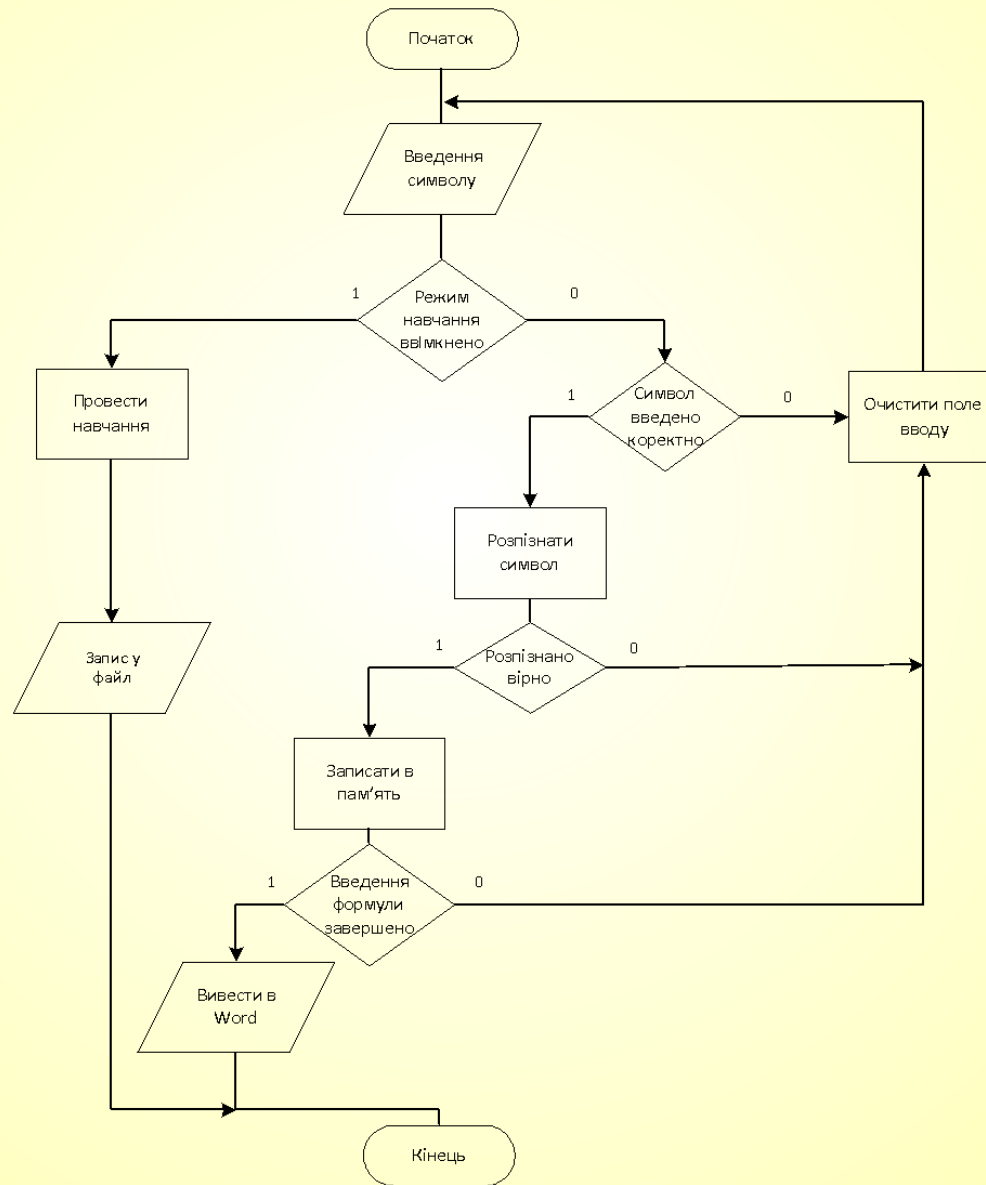
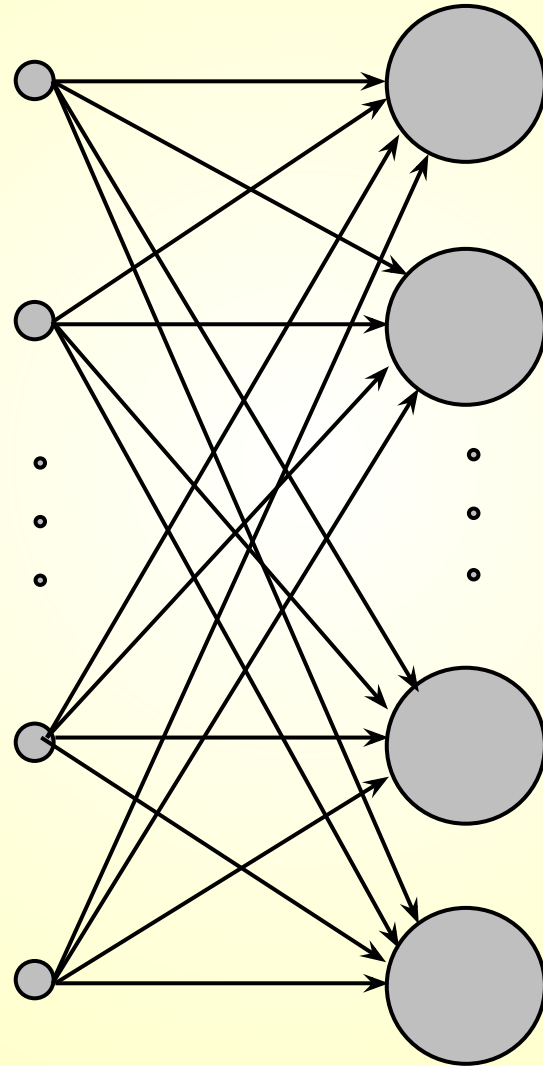


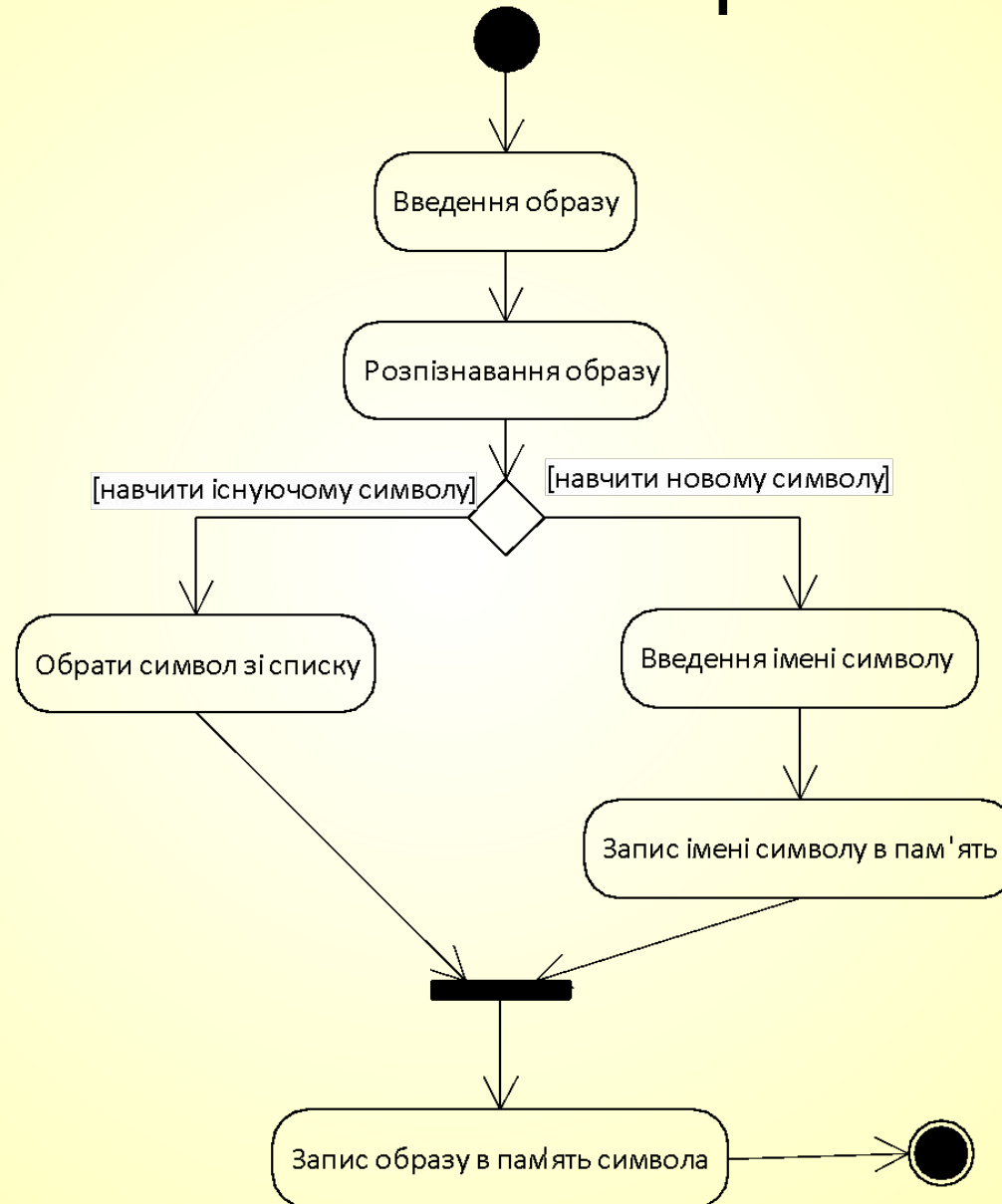
Схема алгоритму роботи модуля



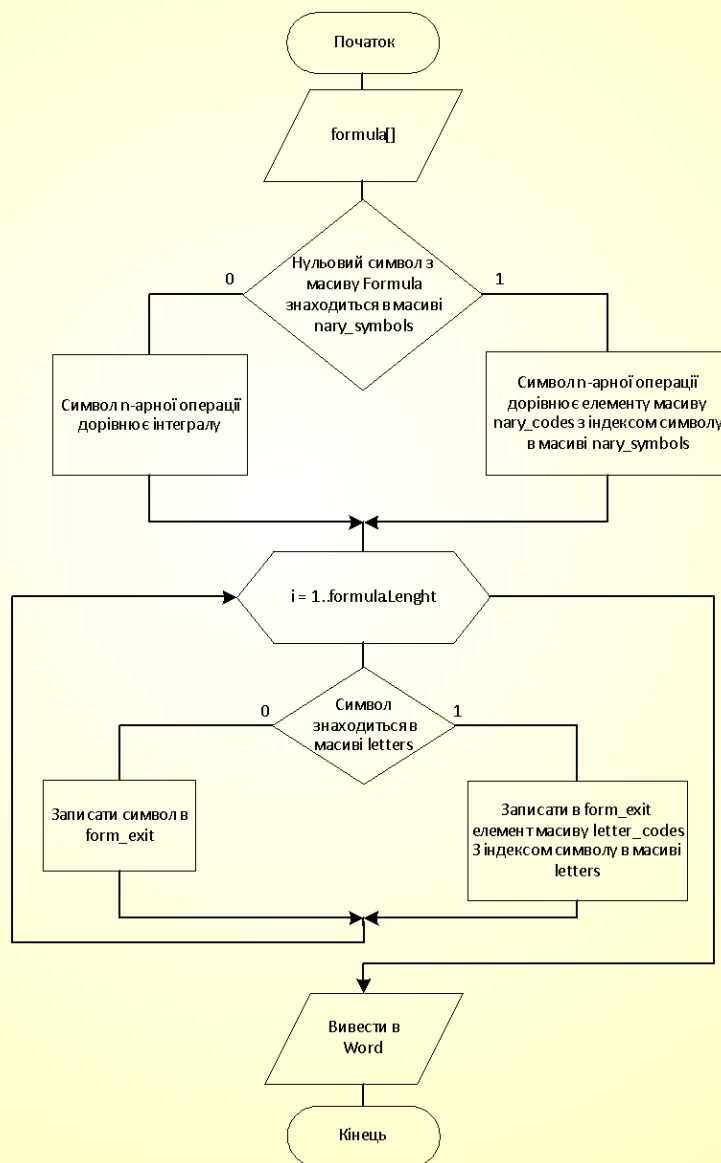
Структура нейронної мережі



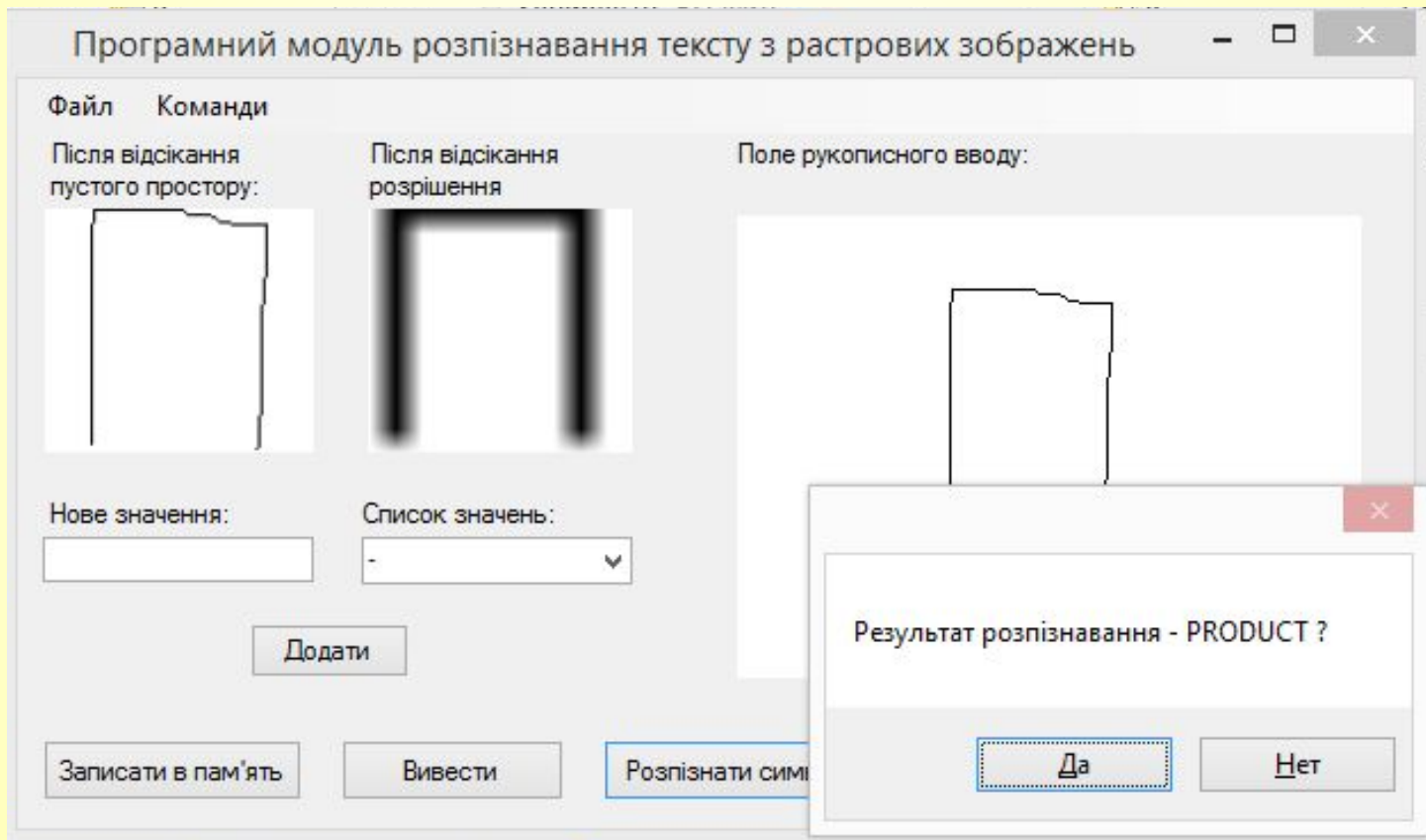
Процедура навчання нейронної мережі



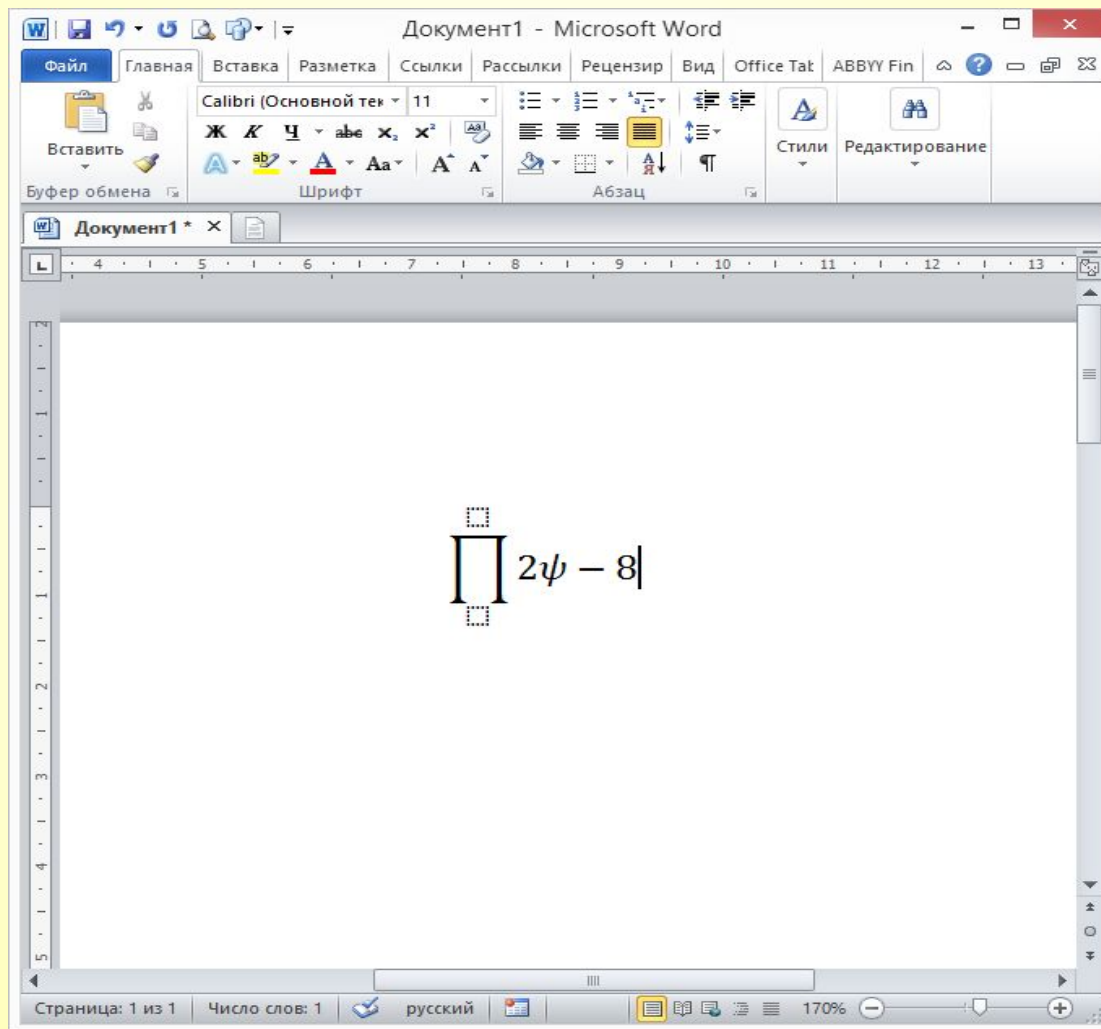
Алгоритм виведення результатів



Приклад роботи програмного модуля



Приклад роботи програмного модуля



Таблиця результатів роботи модуля

Група символів	Відсоток розпізнавання символу		
	Програмний модуль РТРЗ	ABBYY FineReader	OmniPage
Символи грецького алфавіту	85,6	90,1	89,1
Символи латиниці	84,4	93,9	97,2
Арифметичні операції	89,9	98,9	94,5
Цифри	85,3	98,9	98,2
Спеціальні математичні символи	95,1	35,8	33,4
Середнє значення	88,67	83,52	82,48

Висновки

1. Проаналізовано сучасні програмні рішення розпізнавання тексту з растрових зображень та методи розпізнавання тексту з растрових зображень і обрано технологію з використанням нейронної мережі.
2. Розроблено структуру програмного модуля розпізнавання тексту з растрових зображень, яка містить блок розпізнавання на основі нейронної мережі.

Висновки

3. Реалізовано алгоритм розпізнавання спеціальних текстових символів на основі нейронної мережі, що дозволило розпізнавати спеціальні математичні символи.
4. Реалізовано алгоритм передачі спеціальних текстових символів в *MS Office Word*, що дозволило коректно відобразити спеціальні математичні символи в *MS Office Word*.

Висновки

5. Проаналізовано роботу програмного модуля. Середній відсоток розпізнавання символів незалежно від групи становить 88,67%, порівняно з аналогами на ринку OCR систем є в середньому на 5,67 % вищим показником
6. Практична значимість результатів проектування полягає в розпізнаванні рукописного тексту та оцифрування даних, що розміщені на паперових носіях, зокрема математичних формул.

Дякую за увагу