

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ  
ІМЕНІ ТАРАСА ШЕВЧЕНКА**

Факультет комп'ютерних наук та кібернетики  
Кафедра теоретичної кібернетики

**Кваліфікаційна робота**  
на здобуття ступеня бакалавра  
за спеціальністю 122 Комп'ютерні науки  
на тему:

**РОЗРОБКА АЛГОРИТМІВ РОЗПІЗНАВАННЯ РУКОПИСНИХ ЛІТЕР**

Виконала: студентка 4-го курсу  
Анастасія БОЙКО

\_\_\_\_\_ (підпис)

Науковий керівник:  
професор кафедри теоретичної кібернетики  
доктор фіз.-мат. наук, професор  
Юрій КРАК

\_\_\_\_\_ (підпис)

Засвідчую, що в цій роботі  
немає запозичень з праць інших авторів  
без відповідних посилань.

Студент \_\_\_\_\_ (підпис)

Роботу розглянуто й допущено до захисту  
на засіданні кафедри теоретичної кібернетики  
« \_\_\_\_ » \_\_\_\_\_ 2022 р.,  
протокол № \_\_\_\_  
Завідувач кафедри  
доктор фіз.-мат. наук, професор  
Юрій КРАК

(підпис)

## РЕФЕРАТ

Обсяг роботи 40 сторінок, 41 ілюстрація, 3 таблиці, 21 джерело посилань.

Ключові слова: ТЕХНОЛОГІЇ РОЗПІЗНАВАННЯ ТЕКСТУ, РОЗПІЗНАВАННЯ РУКОПИСНОГО ТЕКСТУ, РОЗПІЗНАВАННЯ РУКОПИСНИХ ЛІТЕР, ОПТИЧНІ ТЕХНОЛОГІЇ РОЗПІЗНАВАННЯ ТЕКСТУ

Об'єктом розробки є розпізнавання рукописного тексту.

Метою роботи є аналіз сучасного стану технологій з розпізнавання рукописних літер.

Методи дослідження. При написанні даної роботи використовувалася сукупність методів: спостереження, дедукція і індукція, експеримент, аналогія, ранжування, метод узагальнення, метод абстрагування та інші загальнонаукові методи.

Можливі сфери застосування. Результати даного дослідження можуть використовуватися для подальшого вивчення питання розпізнавання рукописного тексту.

Результати роботи: проведено аналіз методів та алгоритмів що використовуються для розпізнавання рукописного тексту. Встановлено, що розвиток технологій офлайн та онлайн розпізнавання значним чином відрізняється. Що доведено експериментальним шляхом.

## ЗМІСТ

|  |    |
|--|----|
| РЕФЕРАТ  | 2  |
| ВСТУП  | 4  |
| РОЗДІЛ 1. ТЕОРЕТИЧНІ Й ПРАКТИЧНІ ЗАСАДИ ПОБУДОВИ ТЕХНОЛОГІЙ РОЗПІЗНАВАННЯ РУКОПИСНИХ ЛІТЕР | 6  |
| 1.1. Характеристика існуючих технологічних комплексів розпізнавання рукописних літер       | 6  |
| 1.2. Структурні елементи побудови методів оптичного розпізнавання рукописних літер         | 13 |
| РОЗДІЛ 2. АНАЛІЗ АЛГОРИТМІВ РОЗПІЗНАВАННЯ РУКОПИСНИХ ЛІТЕР                                 | 18 |
| 2.1. Основні принципи функціонування алгоритмів розпізнавання рукописних літер             | 18 |
| 2.2. Особливості практичної реалізації комплексів розпізнавання тексту                     | 23 |
| ВИСНОВКИ   | 36 |
| ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ   | 38 |

## ВСТУП

**Оцінка сучасного стану об'єкта розробки.** Технології оптичного розпізнавання тексту є предметом досліджень чисельних вітчизняних й зарубіжних вчених (А. Богуславський [1], Ю. Болотова [2], А. Друки [3], М. Мілешин [3], В. Спіцин [4], Н. Фан [4] та ін.).

**Актуальність роботи та підстави для її виконання.** Розпізнавання текстових елементів є одним з актуальних завдань в силу розвитку інформаційних технологій й все більшого застосування технологій в житті людини. Необхідність фіксування й зберігання документів, що містять рукописні фрагменти, особливо в системах документообігу, зумовлює підвищений інтерес наукового суспільства до цієї проблеми. З іншої сторони, посилення взаємодії людини з технологічними комплексами й програмними продуктами з погляду на зростаючі вимоги до прискорення цієї взаємодії робить питання розпізнавання рукописного тексту машиною актуальним.

**Мета й завдання роботи.** Метою роботи є аналіз сучасного стану технологій з розпізнавання рукописних літер. Для досягнення цієї мети поставлено такі завдання:

- Дослідити існуючі технологічні комплекси розпізнавання рукописних літер та надати характеристику.
- Визначити структурні елементи побудови методів оптичного розпізнавання рукописних літер.
- Дослідити алгоритми розпізнавання рукописних літер.
- Виявити основні принципи функціонування алгоритмів розпізнавання рукописних літер.
- Провести експеримент з метою встановлення ефективності сучасних комплексів з розпізнавання рукописних літер.

**Об'єкт дослідження** – розпізнавання рукописного тексту.

**Предмет дослідження** – методи, алгоритми та підходи до розпізнавання рукописного тексту.

**Методи дослідження.** При написанні даної роботи використовувалася сукупність методів: спостереження, дедукція і індукція, експеримент, аналогія, ранжування, метод узагальнення, метод абстрагування та інші загальнонаукові методи.

**Можливі сфери застосування.** Результати даного дослідження можуть використовуватися для подальшого вивчення питання розпізнавання рукописного тексту.

## РОЗДІЛ 1. ТЕОРЕТИЧНІ Й ПРАКТИЧНІ ЗАСАДИ ПОБУДОВИ ТЕХНОЛОГІЙ РОЗПІЗНАВАННЯ РУКОПИСНИХ ЛІТЕР

### 1.1. Характеристика існуючих технологічних комплексів розпізнавання рукописних літер

Від самого початку розвитку інформаційних технологій стояло гостро питання побудови ефективної взаємодії людини й технологічних комплексів. Ключове місце в цьому процесі відводиться проблемі синхронізації комунікаційних форматів людини й техніки для досягнення певної низки цілей, однією з яких є оптичне розпізнавання тексту. Завдання розпізнавання тексту прийнято вирішувати за допомогою так званих систем оптичного розпізнавання тексту. Системи оптичного розпізнавання тексту являють собою програмні компоненти або цілі технологічні комплекси, що виступають певним посередником-перекладачем між людиною й комп'ютером [5].

Найпершою спробою створення системи оптичного розпізнавання тексту вважають метод Г. Таушека, створений на базі патернів розпізнавання у 1929-1935 рр. Ця система була у подальшому вдосконалена Д. Х. Шепардом за допомогою криптографічних шаблонів у 1950 рр. Через 10 років у результаті співпраці «Readers Digest» й «RCA» була створена система для зчитування серійних номерів на базі комп'ютера RCA 301. У 1974 р. Р. Курцвейл здійснив спробу створення першого комплексу з розпізнавання тексту, що включав сканер й синтезатор мовлення, яка у подальшому стала основою для розвитку компанії «Ксерокс». На початку свого розвитку системи були недосконалими. Однак, сьогодні системи розпізнавання друкованого тексту досягли значного рівня розвитку: 1) точність перевищує 99%, 2) є велика кількість програмних продуктів, як окремих, так й вбудованих в інші програмні комплекси (наприклад, ABBYY FineReader, FreeOCR, OCRopus та ін.), 3)

значна кількість мов (наприклад, NewOCR.com підтримує 29 мов); 4) комерційні й некомерційні; 5) онлайн та офлайн; 6) призначені для друкованого тексту та з можливістю розпізнавання рукописного тексту (HWR - handwriting recognition) [5].

Класично системи розпізнавання поділяють за наступними критеріями: 1) кількість ознак, що використовуються (складні та прості); 2) здатність до навчання (нездатні до навчання, здатні до навчання й ті, що передбачають автоматичну адаптацію (самонавчання)); 3) рівень структурної складності (однорівневі, багаторівневі в залежності від кількості засобів одержання апостеріорної інформації); 4) за характером ознак розпізнавання (детерміновані, ймовірнісні, логічні, структурні (лінгвістичні), нейромережеві, комбіновані) [6]. Структурно класифікація систем розпізнавання наведена на рисунку 1.1.

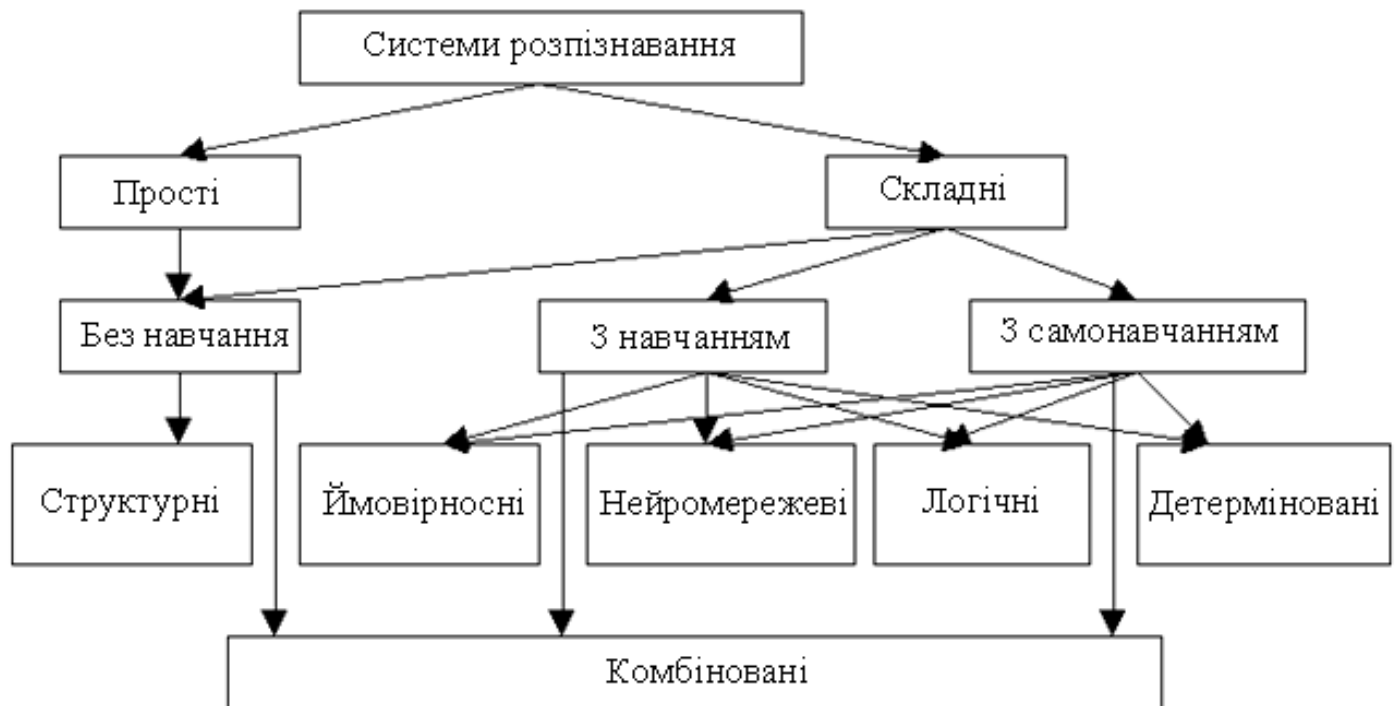


Рисунок 1.1. Класифікація систем розпізнавання тексту [6].

В класифікації окремо не пропонується поділ систем відповідно до типу тексту, оскільки більшість сучасних систем пропонує рішення для обох типів.

Сам процес розпізнавання (англ. optical character recognition, OCR) являє собою механічне або електронне переведення зображень тексту у послідовність кодів, які

«зрозумілі» для електронно-обчислювальної машини. Відокремлюється два основних типи завдань рукописного розпізнавання тексту: 1) онлайн, що передбачає миттєве розпізнавання, тобто розпізнавання тексту відбувається паралельно написанню тексту; 2) офлайн, тобто до моменту самого розпізнавання текст попередньо оброблюється й подається на розпізнавання у синтезованому вигляді. Ключовою відмінністю вказаних завдань є обсяг вхідної інформації. Особливістю паралельного процесу є те, що система може відслідковувати й аналізувати безпосереднім чином створення рукописного тексту, тобто спостерігає за механікою. Таким чином, система отримує не тільки зображення текстового елементу, а й може проаналізувати тиск на інструмент написання, нахил, вектор проведення лінії, швидкість й напрямок руху руки, структуру текстового елементу, порядок утворення текстових фрагментів та ін.

При офлайн розпізнаванні система отримує тільки доступ до зовнішнього вигляду текстових елементів. Формат рукописних текстів характеризується достатньо низькою якістю першочергового зображення, що пов'язано не тільки з індивідуальним характером написання символів, особливостями почерку, а й легкістю з якою псується вихідний документ – забруднення та інші пошкодження паперу призводять до суттєвого зниження якості зображення. У зв'язку з цим, офлайн розпізнавання є більш складним завданням.

Розглянемо відмінності двох завдань. На рисунку 1.2. зображено два близьких за написом слова.



Рисунок 1.2. Схожі за написом слова [7].

Як ми бачимо на зображенні, відмінності є незначні у літерах «о» та «а»: невелика рисочка у другій літері у слові «палка», що відсутня у літері «о» слова

«полка» (рисунок 1.3, а); незначна деформація ліній, що поєднує вказані літери з наступною «л» (рисунок 1.3, б).

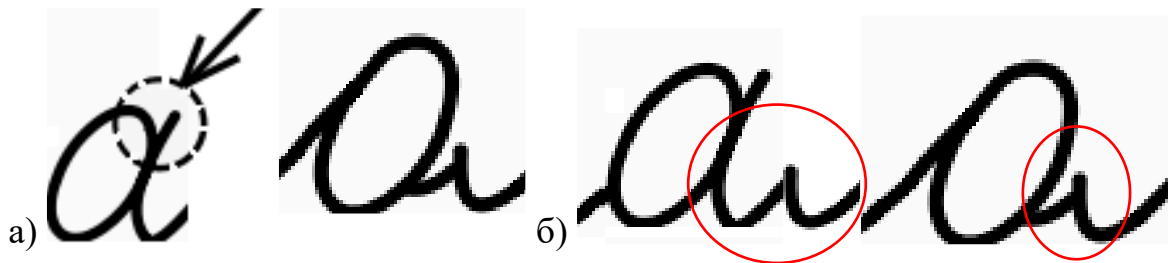


Рисунок 1.3. Відмінності у літерах «а» та «о» [7].

Розглянемо механізм написання цих літер (рисунок 1.4).

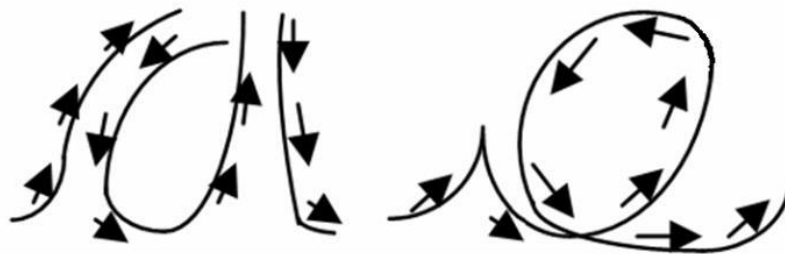


Рисунок 1.4. Механізм написання літер «о» та «а» [7].

Як ми бачимо на рисунку, алгоритми написання літер повністю відрізняються. Кожна літера має власний алгоритм написання, який не залежить від особливостей почерку та ін. Отже, можна зробити висновок, що процес паралельного розпізнавання більш «зрозумілий» для програмних систем й менше залежить від додаткових факторів, що можуть впливати на вихідні дані. Паралельний метод розпізнавання на даний момент широко застосовується у багатьох популярних продуктах, що переважно доступні широкій аудиторії: 1) операційна система Windows, починаючи з 7 версії дозволяє використовувати рукописне введення даних у багатьох програмних компонентах; 2) вузькоспеціалізований продукт «PenReader», що спеціалізується виключно на розпізнаванні рукописного тексту; 3) пошуковий програмний продукт від Google дозволяє вводити дані у перекладач за допомогою спеціальної панелі рукописного введення даних (рисунок 1.5). Як ми бачимо, перекладач є системою, здатною до самонавчання, шляхом взаємодії із користувачем, що має змогу вибрати правильний варіант розпізнаного тексту.

Однак, у паралельного розпізнавання також є певні недоліки, що, передусім пов'язані з технічними вимогами до пристроїв. Найголовніша вимога полягає у тому, щоб користувач був забезпечений екраном достатнього розміру з високою чутливістю до натискання.

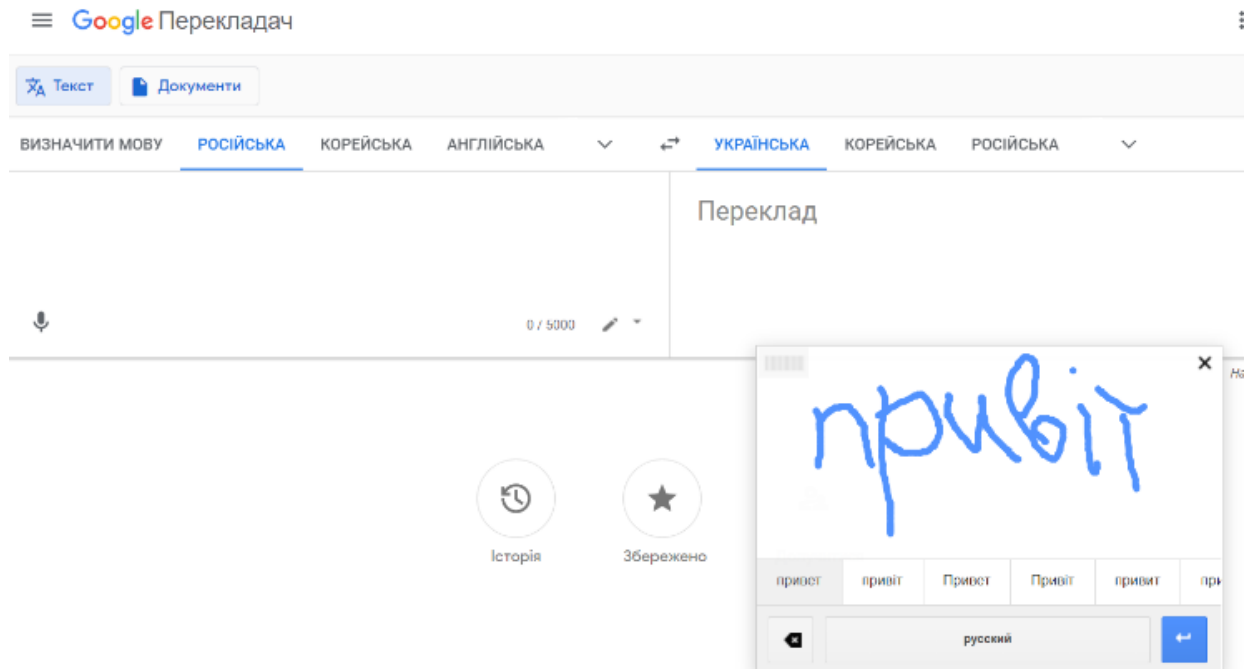


Рисунок 1.5. Реалізація рукописного введення даних у перекладачі Google.

Офлайн розпізнавання тексту базується на зовсім іншому алгоритмі. Спочатку, документ з рукописним текстом проходить процедуру оцифрування. Отримане зображення оброблюється, щоб підвищити якість, яка є принциповою для даного підходу й передається до програмної обробки. На рисунку 1.6. представлена послідовність обробки розпізнавання підпису. Вихідними даними є паперовий носій з рукописним текстом (рисунок 1.6, а). Обов'язковим є переведення в цифровий вигляд (наприклад, за допомогою сканеру чи фотокамери) (рисунок 1.6, б).

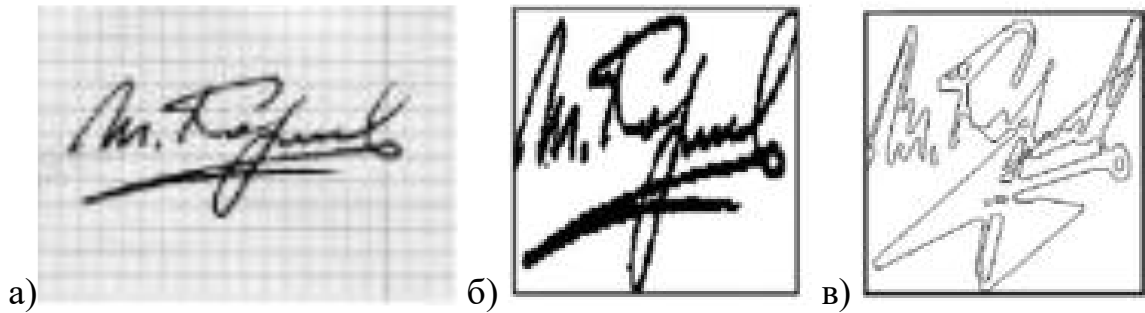


Рисунок 1.6. Послідовність розпізнавання підпису при офлайн підході (а – зразок підпису на папері; б – відсканований підпис; в – попередня обробка графічного зразка, векторне зображення) [8].

На рисунку 1.6, в ми бачимо один з варіантів обробки зображення – отримання растрового варіанту. Цілком очевидно, що даний підхід є найбільш затратним по часу й зусиллям.

Варто звернути увагу на досить популярний продукт, що поєднує в собі обидва підходи - Swype клавіатура.



Рисунок 1.6. Принцип функціонування клавіатури Swype [9].

Сутність даного продукту в тому, що користувач плавно проводить по буквам, зупиняючись на місці пробілу, а пристрій, на якому встановлено таку клавіатуру, вже отримує розпізнаний текст. Нажаль, ще у 2018 р. компанія-власник цього винаходу повідомила про завершення підтримки цього продукту у зв'язку з тим, що більшість

сучасних програмних систем вже автоматично включають можливості для розпізнавання рукописного введення інформації.

## 1.2. Структурні елементи побудови методів оптичного розпізнавання рукописних літер

Розглянемо принципи й основні елементи побудови методів оптичного розпізнавання рукописних літер. На рисунку 1.7 представлена структура оптичного розпізнавання тексту.

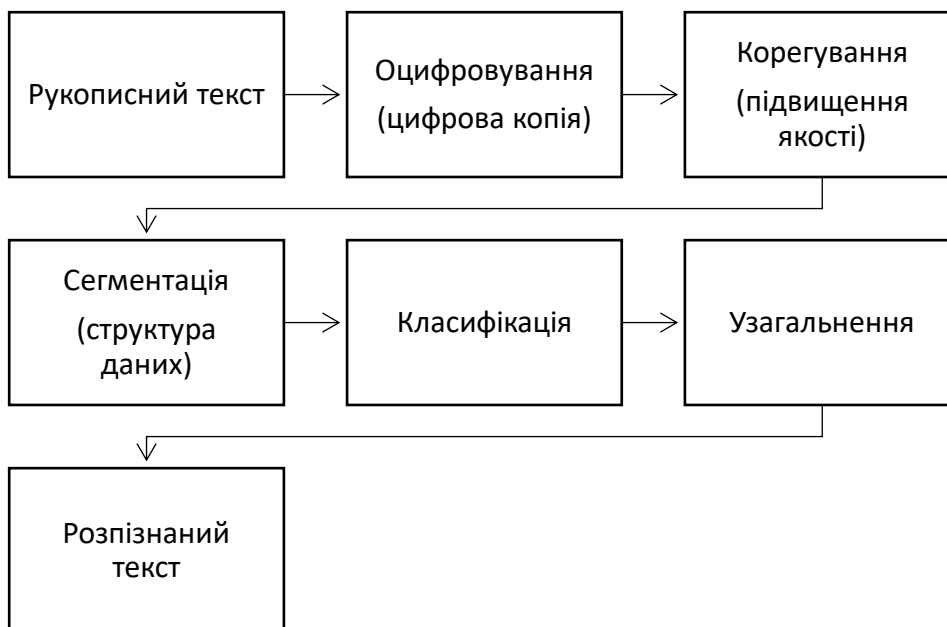


Рисунок 1.7. Структура процесу оптичного розпізнавання тексту [10].

Відокремлюють три принципи, на яких побудовані системи оптичного розпізнавання тексту: 1) принцип цілісності образу, 2) принцип цілеспрямованості, 3) принцип адаптивності [11].

Після отримання цифрової копії рукописного зображення відбувається обробка, головне завдання якої максимально підвищити якість зображення, яка для цього методу є критичною. На цьому етапі видаляються оптичні перешкоди, корегується яскравість, контрастність та інші параметри зображення, щоб отримати максимально чітке зображення. Найрозповсюдженими методами оптимізації зображення є: метод адаптивної фільтрації (використання лінійного усереднення пікселей відносно сусідніх пікселей), метод медіанної фільтрації, математичної морфології, метод

розмиття за Гаусом, методи на основі дискретних вейвлет-перетворень, метод головної компоненти, анізотропної дифузії, фільтр Вінера [12].

Потім відбувається процес сегментації зображення, завдання якого полягає у пошуці й відокремленні структурних елементів тексту (рядки, блоки, символи та ін.). В результаті сегментації ми отримуємо структуру даних (від більшого до меншого): сторінка, що складається з блоків – блок, що складається з рядків – рядок, що складається зі слів – слово, що складаються з літер – літера. Доречним є застосування додаткової корекції на цьому етапі, що дозволить підвищити якість розпізнавання символів за допомогою наступних методів: аналіз сполучень літер, що є характерними для певної мови; словник, граматичний аналіз та ін.

Наступним етапом є класифікація. Найбільш вживаними є наступні типи класифікаторів: шаблонний (растровий); ознаковий; структурний. Шаблонний реалізується шляхом порівняння даних з шаблонами, що вже існують у базі даних. Найпростіші – крапкові, що базуються на тому, що графічні елементи являють собою матрицю з крапок (рисунок 1.8). На жаль, стверджувати, що це саме літера «л» без контексту неможливо з 100% впевненістю.

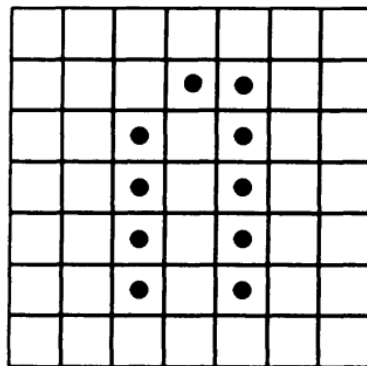


Рисунок 1.8. Зразок крапкового шейпу літери «л» [13].

Структурно процес шаблонної класифікації має наступний вигляд: 1) шейп; 2) порівняння з еталоном, що включає пошук по базі шаблонів й класифікацію за критерієм порівняння; 3) узгодження (результат розпізнавання) [13].

На окрему увагу заслуговує структурно-плямовий еталон й фонтанна (від англ. font – шрифт) реалізація. Цей алгоритм включає функції шаблонного й структурного методів. При розпізнаванні увага приділяється ключовим крапкам об'єкту (плямам): кінці лінії, вузли ліній, місце перехрещування ліній та ін. Сукупність плям складає граф, що й виступає об'єктом пошуку [14].

Основою однакових класифікаторів є порівняння за певними наборами ознак, тобто розпізнавання тексту відбувається не за символом, а за набором ознак, що є властивими цим символам.

Сутність функціонування структурного класифікатора полягає у трансформації шейпів у топологічні уявлення, що віддзеркалюють інформацію про розташування змістовних елементів. Цей метод є найбільш розповсюдженим на даний момент.

Класифікатор використовує один з наступних трьох методів: шрифтовий (multifont), що також зустрічається під назвою символний, шаблонний; нешрифтовий, відомий також як алгоритмічний, інтелектуальний та ознаковий (omnifont), адаптаційний [15].

Структура шрифтового методу передбачає наступні етапи: 1) виділення блоків з текстом; 2) ідентифікація шрифту, що використовується у кожному з блоків; 3) відокремлення символів; 4) аналіз та ідентифікація символів окремо; 5) повторна ідентифікація символів, які не вдалося розпізнати шляхом застосування граматичних правил. Ключовим етапом є ідентифікація шрифту, оскільки неможливість виявити аналог автоматично робить завдання неможливим. Цей фактор обумовлює неможливість застосування вказаного методу для розпізнавання рукописного тексту [15].

В центрі нешрифтового методу – ідентифікація символів за нечіткими атрибутами: наявність / відсутність круглих елементів; тип лінії (вертикальна, горизонтальна тощо), особливість використання простору та ін., що реалізується за допомогою цілої низки алгоритмів або комплексів (топологічні, нейронні та ін.) [15].

Адаптивний метод включає переваги як шрифтового, так й нешрифтового методів, процес застосування якого можна представити наступним чином: 1) формування бази символів за допомогою нешрифтового методу; 2) формування шрифту; 3) застосування шрифтових алгоритмів. Цей метод також не є 100% ефективним для рукописного тексту, однак, дозволяє виявити певні особливості. Для підвищення даного методу необхідна значна база зразків.

Очевидно, що розглянуті методи побудовані на одному принципі – ідентифікація окремих символів, що, з урахуванням особливостей рукописного тексту, робить їх неефективними для розпізнавання.

Найбільш ефективними на даний момент є алгоритми, здатні до самонавчання, в тому числі й за допомогою нейронних мереж. Основою нейронних мережевих алгоритмів є модель нейрона, що включає три типи елементів: 1) набір синапсів з урахуванням його ваги; 2) сума; 3) активація (рисунок 1.9). Як ми бачимо на рисунку 1.9 нейронна мережа ідентифікує окремі елементи й шляхом співставлення проводить ідентифікацію символу.

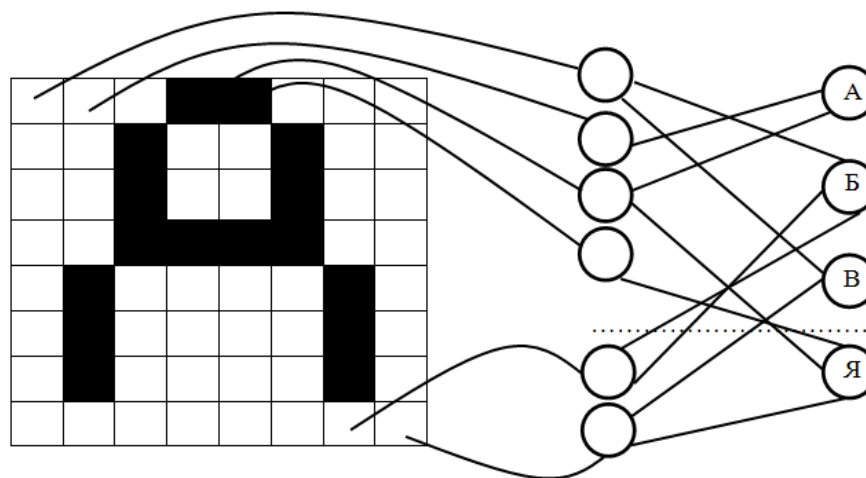


Рисунок 1.9. Графічне зображення принципу роботи нейронної мережі [11].

Перевагою застосування нейронних систем є здатність до самонавчання. При рукописному тексті варіативність занадто висока. Щоб знизити вплив даного фактору використовується примітивізація (розклад комплексу на мінімальні змістовні). Незважаючи на популярність даного підходу, експериментальна частина незначна. З

однієї сторони, рукописний текст можна розділити на лінії, відрізки, фрагменти ліній, що робить цей метод теоретично дійовим. Однак, при рукописному введенні тексту важливу роль відіграє напрямок руху, який неможливо відслідкувати у цифрових відрізках.

Інший підхід, так званий «графовий», передбачає перетворення символу на граф й використання генетичного алгоритму для розширення кола зразків для порівняння. Даний метод дозволив не зважати на напрямок руху при письмі, а зосередити увагу на характерні відмінності (дидактичні знаки). Експеримент з арабською мовою дав достатньо цікаві результати – ідентифікація символів таким чином з застосуванням простих методів розпізнавання продемонстрували результат у 80% точності [10].

Таким чином, на підставі аналізу можливостей застосування методів розпізнавання рукописного тексту можна зробити наступний висновок: 1) сьогодні очевидним є прогресивний розвиток систем розпізнавання рукописного тексту; 2) окремі методи не дозволяють отримати високу ефективність; 3) комплекси методів є правильним рішенням; 4) специфіка візуалізації й створення рукописного тексту вимагає зосереджувати увагу не тільки на методах розпізнавання, а і на методах адаптації зображення до вимог аналітичних систем; 5) у процесі розпізнавання друкованого тексту ключовим є шрифтове зображення символу, а у рукописному – структурні змістовні побудови символу.

## РОЗДІЛ 2. АНАЛІЗ АЛГОРИТМІВ РОЗПІЗНАВАННЯ РУКОПИСНИХ ЛІТЕР

### 2.1. Основні принципи функціонування алгоритмів розпізнавання рукописних літер

Значну роль в процесі розпізнавання не тільки рукописних символів, а й будь яких елементів грає попередній аналіз об'єктів з метою виявлення характерних ознак. Отриманий набір ознак дозволяє охарактеризувати об'єкт. Розглянемо підходи для вирішення цієї задачі.

Одним з розповсюджених методів є застосування прихованих марковських моделей, а також поєднання цього методу з нейронними мережами. Даний метод передбачає отримання символів із зображення шляхом розділення слова на вектори, з яких потім формується послідовність ознак марковського ланцюга, що спостерігається (рисунок 2.1) [11].

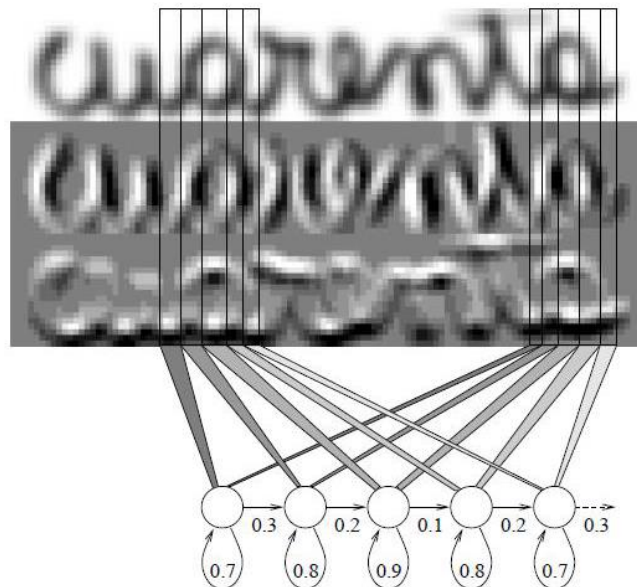


Рисунок 2.1. Застосування методу прихованих марковських моделей для розпізнавання рукописного тексту [11].

Вказаний метод дав можливість отримати точність до 95% при розпізнаванні числівників, написаних словами.

Метод отримання ознак на основі аналізу восьми сусідів базується на сегментуванні зображення на вісім окремих зон, як наведено на рисунку 2.2.

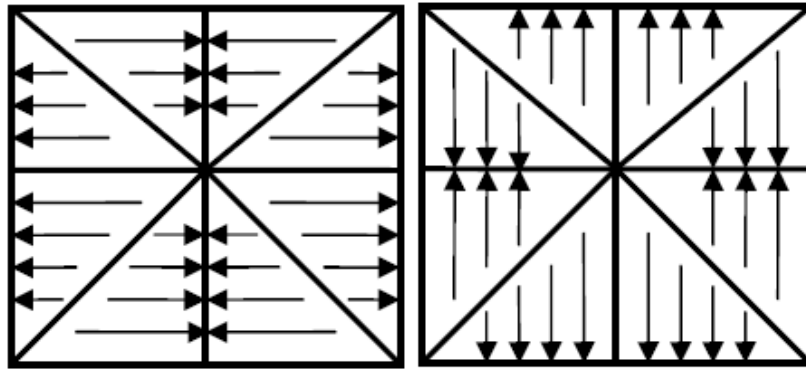


Рисунок 2.2. Метод восьми напрямків [13].

Даний метод довів свою ефективність при використанні для коротких слів, однак, для асиметричних, занадто довгих слів не використовується.

Значною популярністю користується метод отримання ознак, що заснований на розрахунку рівня середньої яскравості (рисунок 2.3). Достатньо ефективний метод для ідентифікації символів невеликого розміру й цифр. Аналогічно попередньому методу не є ефективним для довгих слів [14].

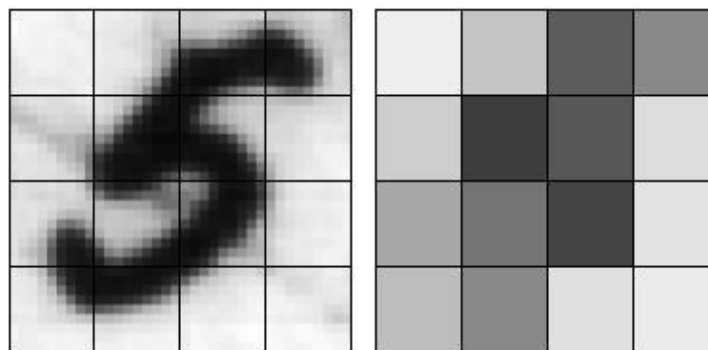


Рисунок 2.3. Застосування градієнтного методу [14].

Використання комплексів методів достатньо ефективно, так для розпізнавання рукописних слів було використано не тільки приховані марковських моделі, а й принцип отримання локальних екстремумів висот елементів відносно базової лінії (рисунок 2.4) [15].

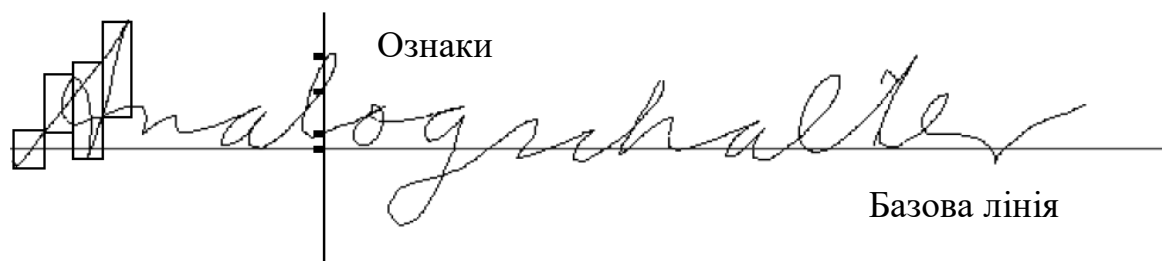


Рисунок 2.4. Метод отримання ознак шляхом порівняння екстремумів висот відносно базової лінії [15].

Комбінації семантичних та статистичних ознак були описані в роботі [10]. При цьому статистичні ознаки – вертикальні й горизонтальні проекції, а сама літера відображається у фіксованому розмірі з використанням бікубічної інтерполяції. На жаль, результати вказаного експерименту довели, що застосування цього методу можливе лише для окремих символів [10].

Комбінований підхід з використанням первинних і вторинних ознак застосовується в роботі [11]. Зображення розбивається на вертикальні зони: нижню, середню і верхню. Аналізується наявність прямих, кіл, верхніх і нижніх штрихів і ряд інших ознак.

Комбінації описаних підходів описані в роботі [12], отримання ознак методом векторизації показано в [13], а так само кілька оригінальних варіантів описано в роботах [14, 15]. Комбінування підходів в кілька разів знизило швидкість обробки даних, а ефективність розпізнавання зросла незначно - на кілька одиниць відсотків.

Результати проведеного мною аналізу по визначенню ефективності методик оптичного розпізнавання рукописного тексту наведено у табл. 2.1.

Найбільш ефективним й перспективним є методи засновані на прихованих марковських моделях, що не тільки показав високу точність, а й, найголовніше, довів ефективність при розпізнаванні довгих слів, написаних разом. Вказаний метод застосовувався тільки для рукописного тексту, написаного іспанською мовою, однак, можна зробити припущення, що аналогічні результати можна отримати й для інших європейських мов.

Табл. 2.1.

Порівняння ефективності методів оптичного розпізнавання рукописного тексту.

| <b>Метод оптичного розпізнавання</b> | <b>Відсоток точності розпізнавання</b> | <b>Коментарі</b>                 |
|--------------------------------------|--|----------------------------------|
| Нейронні мережі                      | До 100%                                | Тільки для окремих символів      |
| Класифікація графічних примітивів    | До 95%                                 | Тільки для окремих символів      |
| Застосування графів                  | До 99%                                 | Тільки для окремих символів      |
| Аналіз окремих сегментів             | До 100%                                | Тільки для окремих символів      |
| Метод опорних векторів               | відсутня                               | Тільки для окремих цифр          |
| Приховані марковські моделі          | До 95%                                 | Повноцінні слова іспанської мови |

У табл. 2.2 наведено результати порівняння методів вилучення символів з зображення.

Табл. 2.2.

Порівняння методів вилучення символів.

| <b>Метод</b> | <b>Недоліки</b>  |
|--------------|--|
| Матричні     | Можливість застосування тільки для окремих символів                            |
| Векторні     | Зниження проценту розпізнавання при збільшенні бази                            |
| Зонування    | Неможливо застосувати для слів з варіативною довжиною; низький рівень чіткості |
| Примітиви    | Недостатній рівень розробки на даний момент                                    |
| Екстремуми   | Низький відсоток розпізнавання   |

Як ми бачимо, на жаль, аналіз існуючих методик відокремлення символів з зображення показав відсутність 100% ефективного метода.

## 2.2. Особливості практичної реалізації комплексів розпізнавання тексту

Для проведення практичного експерименту були обрані наступні програмні продукти, доступні онлайн: Convertio [18] й відомий програмний продукт Abbyu FineReader [17] – обидва продукти обрані для офлайн версії рукописного тексту; Google перекладач – для онлайн, тобто паралельного розпізнавання.

Порівняння властивостей програмних продуктів наведена у табл. 2.3.

Табл. 2.3.

Порівняння властивостей програмних продуктів обраних для експерименту з розпізнавання рукописного тексту.

|  | Convertio                        | Abbyu FineReader  |
|--|----------------------------------|---|
| Доступність онлайн   | +                                | Тимчасово. У подальшому онлайн версія продукту буде анульована                  |
| Доступна кількість сторінок для безоплатного розпізнавання | 10                               | 5   |
| Вартість сторінки при комерційному використанні            | 8 доларів США / 100 сторінок     | 129 євро на рік з можливістю розпізнавати до 5000 сторінок                      |
| Мобільна версія  | -                                | +   |
| Кількість мов  | 68, в тому числі мови країн Азії | 193, в т.ч. мови країн Азії   |
| Формат вихідного файлу                                     | Word, Pdf, Excel, Txt            | Word, Pdf, Excel, Txt, PPT, Open Document, Fiction book, Electronic Publication |

Інтерфейси програм доступні на наступних мовах: Abbyy FineReader (англійська, російська, японська, французька), Convertio (18 мов, в т. ч. європейські та азіатські мови). На жаль, жодний програмний продукт не підтримує українську мову інтерфейсу.

Для зразку офлайн розпізнавання було створено два написи українською (Я люблю Україну) та англійською мовою (I love Ukraine) й зроблено дві фотографії: одна з використанням спалаху, а інша без використання спалаху, щоб ускладнити процес розпізнавання (рисунок 2.5 а, рисунок 2.5 б).

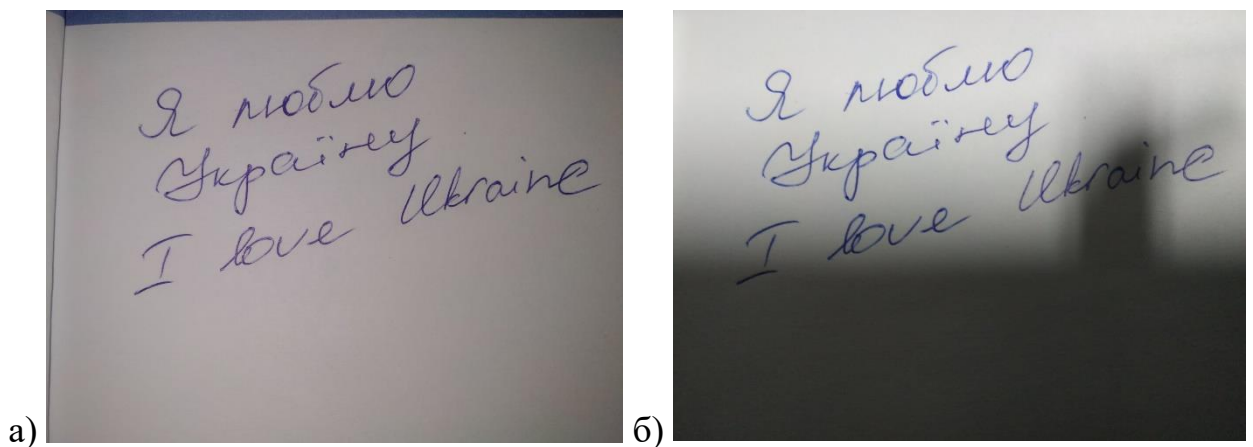


Рисунок 2.5. Зразки для тестування програмних комплексів з розпізнавання рукописного тексту.

Проаналізуємо інтерфейси програм (рисунок 2.6).

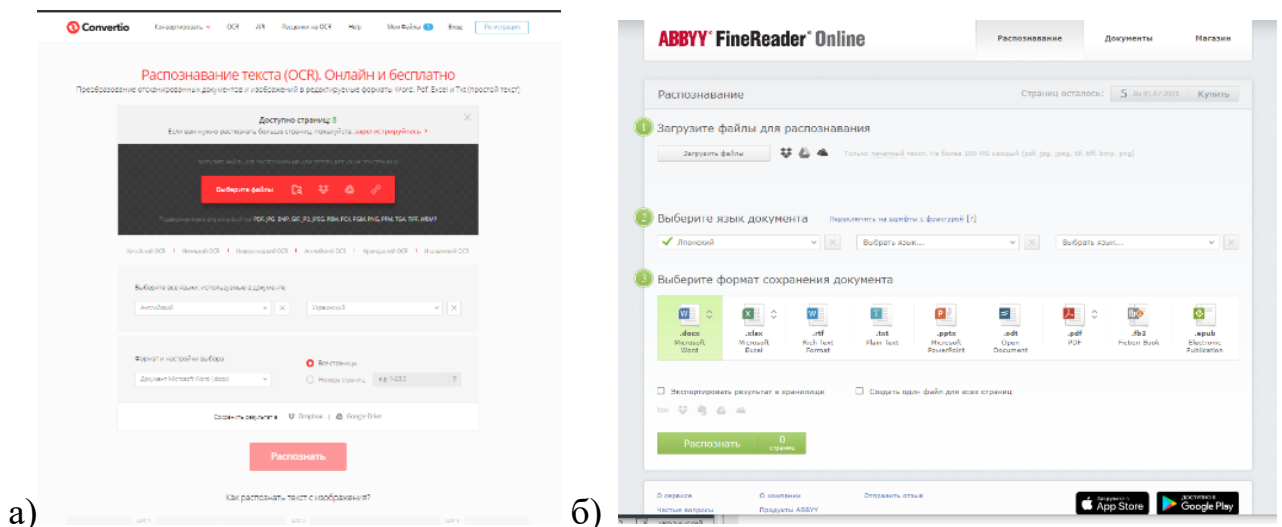


Рисунок 2.6. Інтерфейси програм Convertio (а); Abbyy FineReader (б).

Інтерфейси програм достатньо прості й інтуїтивно зрозумілі. У Convertio місце, куди треба відправити документ виділяється більше, ніж у Abbyy FineReader. У Abbyy FineReader процес розпізнавання представлений ієрархічно у вигляді 3 кроків: 1) завантаження файлу 2) вибір мови документу; 3) вибір формату збереження файлу. У Convertio процес розпізнавання пояснюється під віконцем з розпізнавання, де не тільки зазначені кроки, але й наведено пояснення для користувача у якій послідовності та які саме дії необхідно виконати, щоб отримати готовий файл. Перевагою Convertio є можливість розпізнавати не весь файл, а обрати сторінки які потрібні.

Після того, як зображення були завантажені, ми отримали наступні результати:

1) На платформі Convertio розпізнавання відбулося. Ми отримали файли. Однак, зображення не було розпізнано (рисунок 2.7).

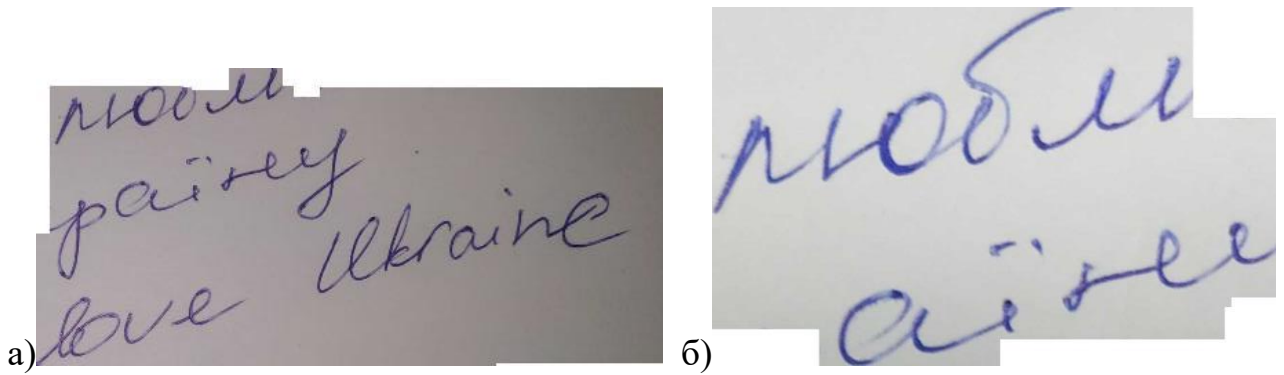


Рисунок 2.7. Отримані результати з розпізнавання експериментальних зображень програмним комплексом Convertio.

На рисунку 2.7 (а) ми бачимо результат розпізнавання зображення, що було навмисно пошкоджене. Програма вичленила фрагмент, однак, важко зрозуміти за яким принципом. Друге зображення (рисунок 2.7 (б)) також показує обрізаний фрагмент з елементами обробки зображення – яскравість отриманого зображення значно вища й текст став чіткішим. Однак, розпізнавання тексту не відбулося.

2) Платформа Abbyy FineReader видала помилку: файл не було оброблено, документ розпізнано з помилками, сторінки не списані (рисунок 2.8).

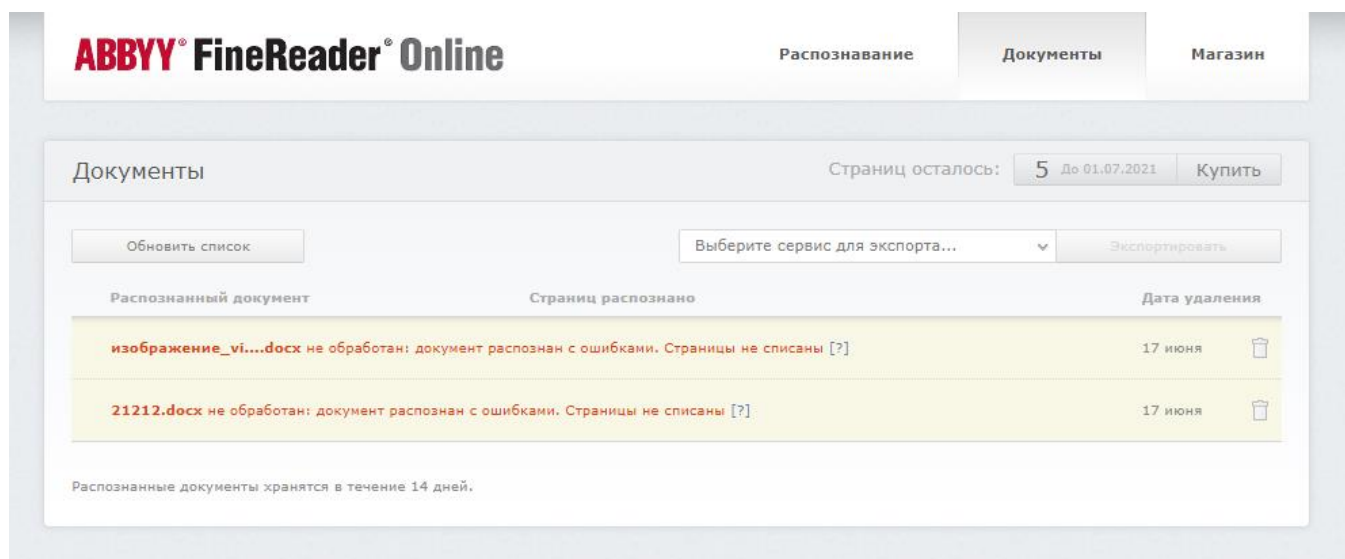


Рисунок 2.8. Помилка у програмі Abbyy FineReader.

Я спробувала ще один сервіс img2txt й отримала цікаві результати. Сервіс також не провів розпізнавання, мотивуючи це тим, що: «Сервер детектував розпізнавання капчі, використовуйте сервіс за призначенням, а не в злісних цілях...» (рисунок 2.9) [19].

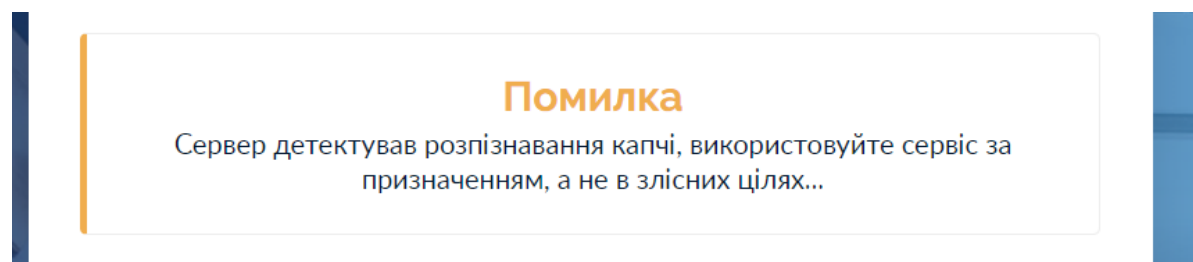


Рисунок 2.9. Помилка на сайті програмного комплексу img2txt.

Таким чином, можна зробити висновок, що офлайн технології не достатньо розвинуті на даний момент.

Експеримент з онлайн розпізнаванням за допомогою рукописного вводу у перекладачі Google дозволив отримати цікаві результати. По-перше, перекладач може ідентифікувати написаний текст за мовою. Спочатку я обрала режим «визначити мову» й почала вводити текст «Я люблю Україну» (рисунок 2.10). Як вже згадувалося у Розділі 1, розпізнавання рукописного тексту на платформі відбувається з урахуванням здатності програмного модулю до самонавчання й взаємодії користувача

безпосередньо з платформою в процесі розпізнавання тексту. Так ми бачимо на рисунку 2.10, що платформа пропонує нам обрати найкращий варіант, з тих, що вона ідентифікувала.



Рисунок 2.10. Фрагмент зони рукописного вводу у перекладачі.  
При введенні повної фрази перекладач розпізнав текст вірно (рисунок 2.11).

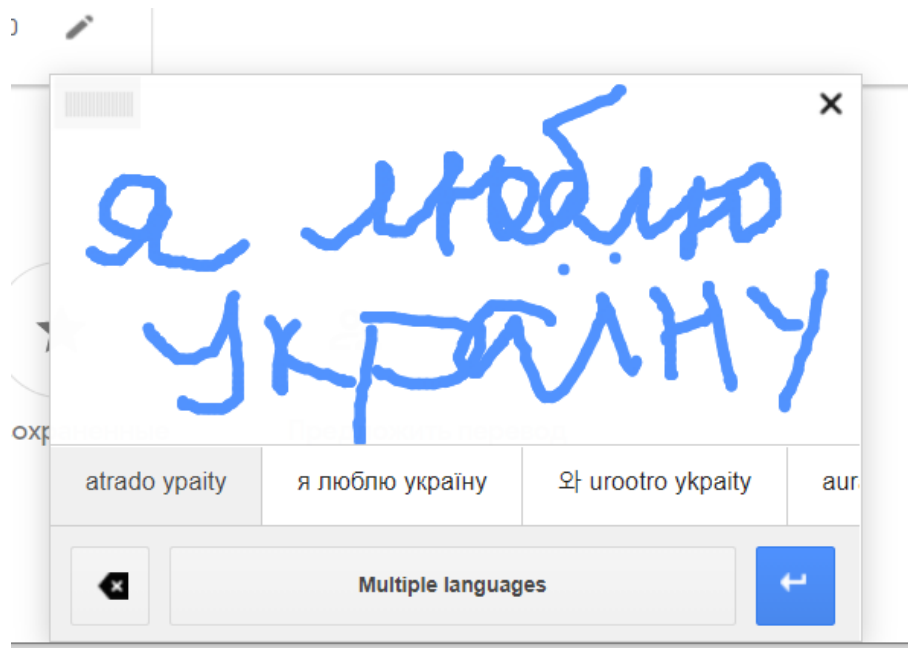


Рисунок 2.11. Розпізнавання повної фрази «Я люблю Україну».

Варто звернути увагу на те, що програмою не враховано регістр – всі літери розпізнані маленькими. Очевидним є застосування бази, оскільки програма надає декілька версій фраз.

Експеримент був проведений ще раз, після обрання мови розпізнавання (рисунок 2.12).

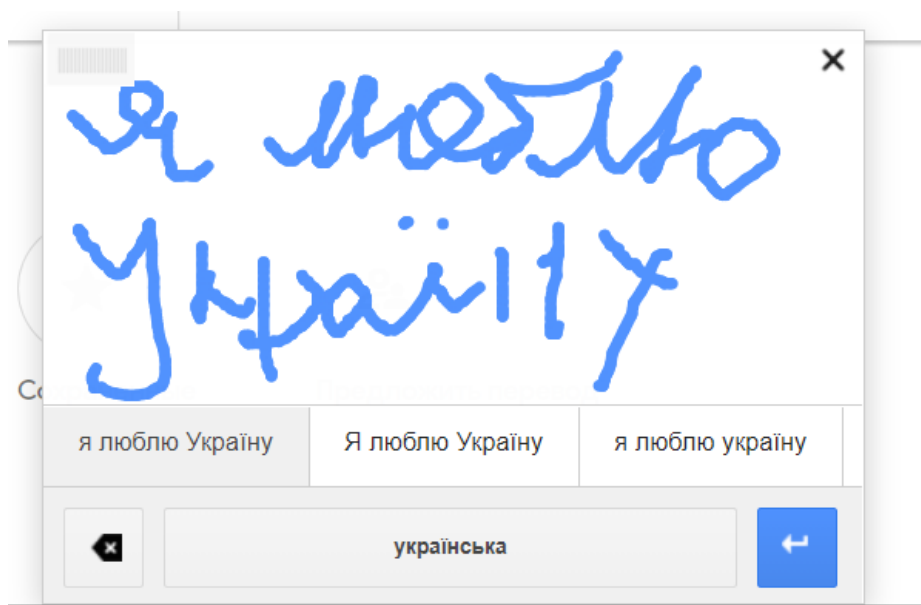


Рисунок 2.12. Експеримент з рукописним розпізнаванням з визначеною мовою тексту.

Текст навмисно був написаний нечітко, намагаючись ускладнити розпізнавання. Однак, переклад впорався з 100% чіткістю. Й, варто звернути увагу, що у запропонованих варіантах враховані особливості мови – наявність великих літер.

Проаналізувавши можливості розпізнавання фрагментів рукописного тексту, розглянемо особливості розпізнавання окремих літер з урахуванням характеристик українськомовного письма. Відповідно до програми навчання грамоти, сформованої відповідно до Державного стандарту початкової освіти на основі вимог Закону «Про освіту» та Концепції Нової української школи, методика навчання української мови обов’язково включає каліграфію й спрямована на засвоєння учнями знань про мову, формування мовних і мовленнєвих умінь та навичок; є наукою про технологію застосування принципів, методів, прийомів і засобів навчання залежно від

дидактичної мети й завдань, характеру матеріалу, етапу навчання, методи формування каліграфічних умінь [20]. Тобто, мова йде про наявність певних стандартів рукописного письма, універсального для кожного, хто вивчає українську мову. Зразок стандартизованого рукописного письма наведений на рисунку 2.13. Варто звернути увагу, що навчальний рукописний шрифт передбачає безвідривне письмо, спрямоване на підвищення його швидкості, а для процесу розпізнавання тексту, відповідно, - додаткове ускладнення завдання.

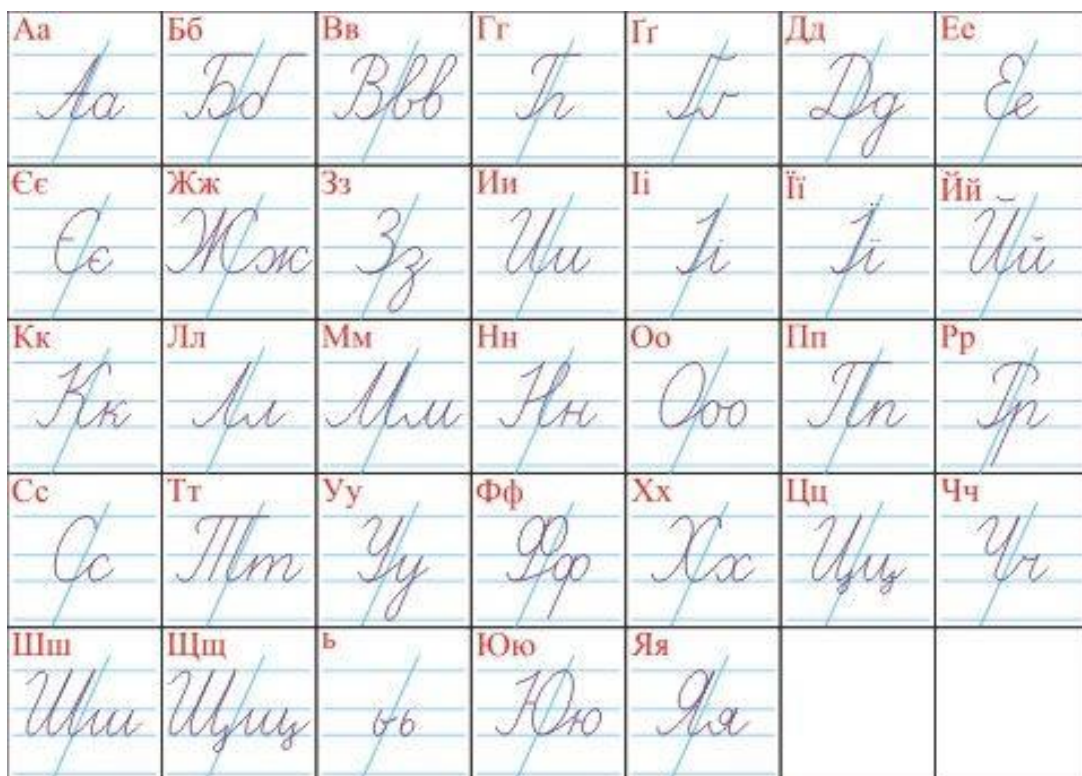


Рисунок 2. 13. Зразок стандартизованого рукописного алфавіту.

У якості матеріалу для експерименту були обрані фрагменти прописів з української мови з найбільш складними за написанням літерами (ж, і, ї, у, б, в та ін.) (рисунок 2.14) [21].

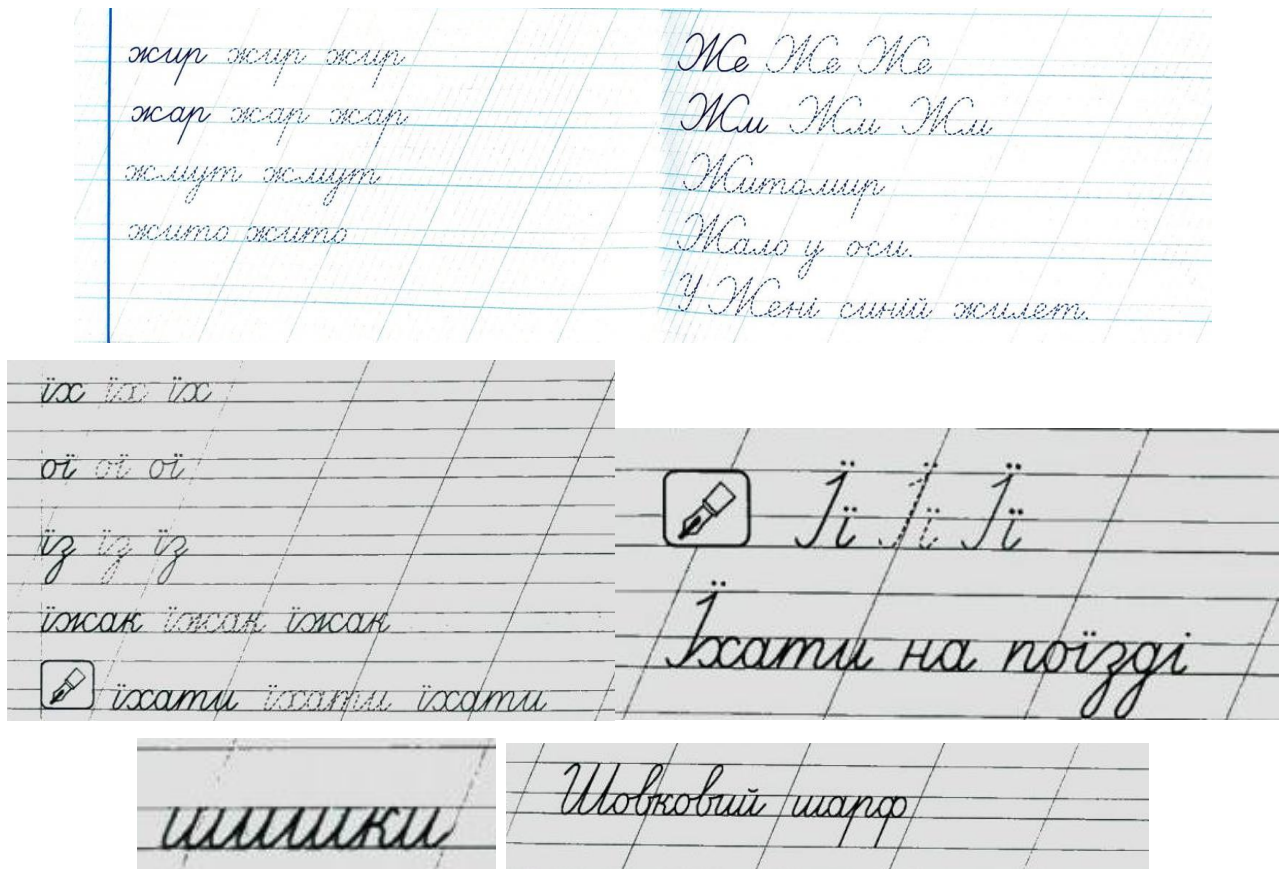


Рисунок 2.14. Зразки для експерименту з розпізнавання стандартизованого рукописного тексту.

Компанія Google пропонує додатковий цікавий інструмент щодо розпізнавання тексту – мобільний додаток, що дозволяє розпізнавати текст, розташований на зображеннях. Інтерфейс програми зручний й пропонує аналогічні функції перекладача на комп’ютері з додатковими опціями роботи з зображенням. На рисунку 2.14 (а) стандартний інтерфейс, що ми бачимо при вході у додаток. У правому нижньому куті є кнопка «камера», при натисканні якої з’являється екран (рисунок 2.14, б), де можна обрати три опції введення зображення: фотографія (“instant”, сканування “scan”, завантаження з галереї “import”).

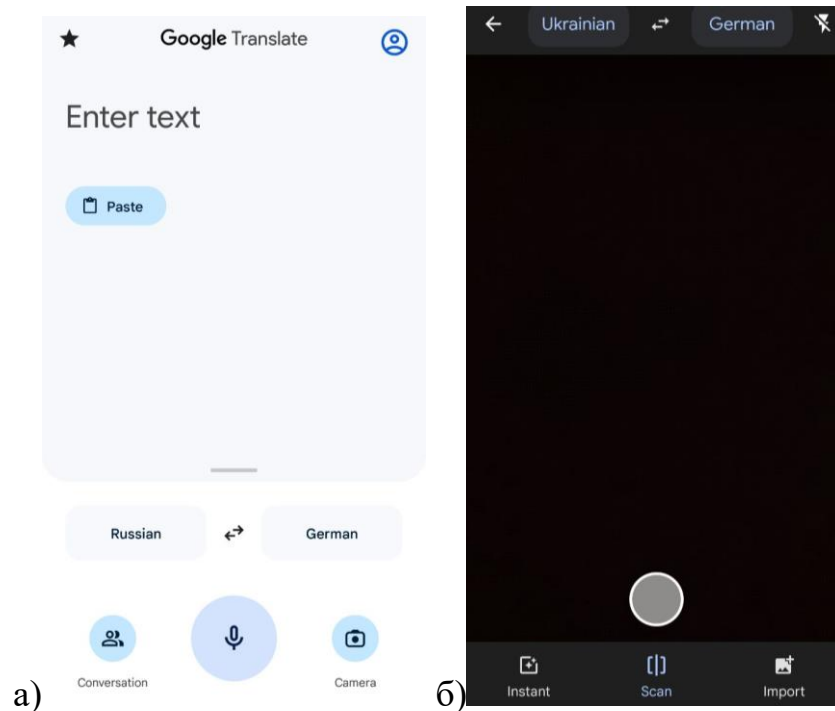


Рисунок 2.15. Інтерфейс мобільного додатку від компанії Google.

Результати експерименту наведені на рисунку 2.16. По-перше, я завантажила зразок за допомогою імпорту з галереї й обрала функцію «визначити мову». Мобільний додаток запропонував обрати частину тексту, що підлягає розпізнаванню. Однак, не зміг визначити ні мову, ні сам текст (рисунок 2.16, а). Я повторила експеримент, обравши українську мову. Однак, програма не знайшла український текст. Отже, можна зробити висновок, що розпізнавання рукописного тексту з зображення у мобільному додатку від компанії Google неможливе.

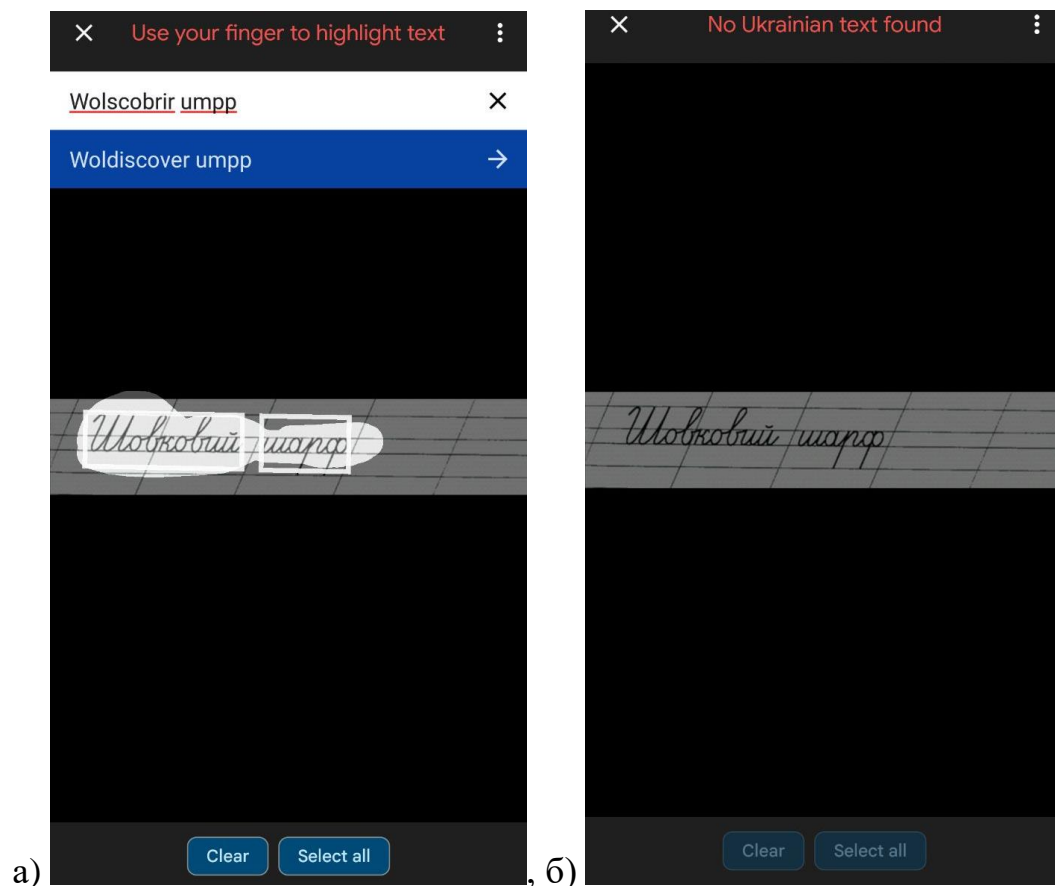


Рисунок 2.16. Експеримент з розпізнавання рукописного, стандартизованого тексту за допомогою мобільного додатку від компанії Google.

Повторимо експеримент, безпосередньо вводячи обрані тексти (жир, жар, жмут, Житомир, Жало у оси, У Жені синій жилет, їхати, Їхати на поїзді, шишки, Шовковий шарф) з наявними маленькими та великими літерами.

Результати експерименту наведені на рисунку 2.17. Слово «жир» було розпізнано коректно (а). Слово «Житомир» було навмисно написане, не дотримуючись стандартизованого письма, й, як ми бачимо, сервіс не зміг розпізнати (б). «Жало у оси» було написано з відхиленнями від стандартів, однак, розпізнавання пройшло вдало (в). Отже, можна зробити висновок, що на якість розпізнавання впливає наявність однакових елементів у алгоритмі написання різних літер (літери «и», «м»). Речення «У Жені синій жилет» дало декілька результатів: 1)

словосполучення «у Жені» не було розпізнано, б) повне речення сервіс ідентифікував з помилкою у слові «жилет» (г, д).

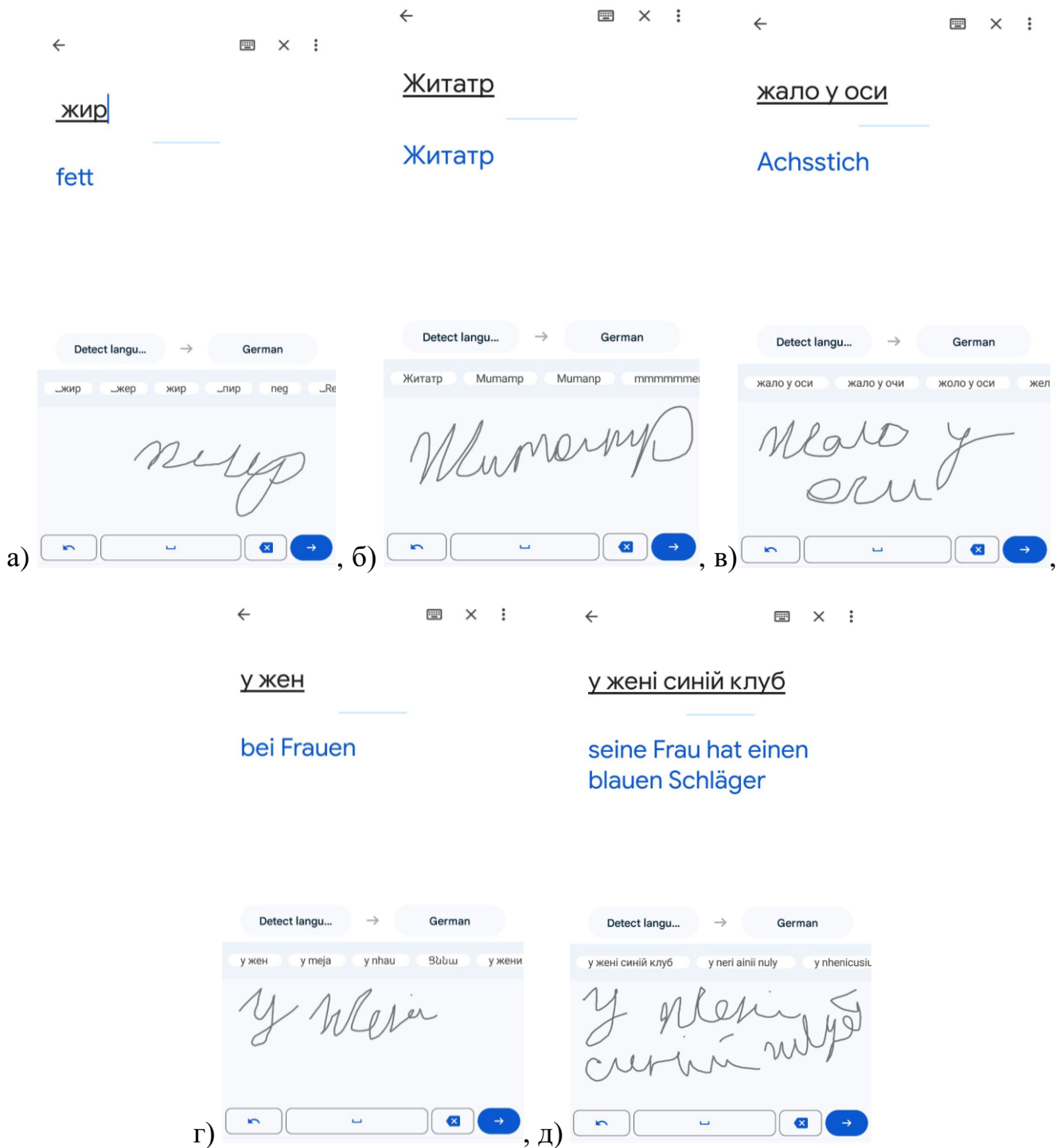


Рисунок 2.17. Результати експерименту з розпізнавання рукописного тексту за допомогою мобільного додатку Google (літери «ж», «е», «н», «і», «и», «с», «н», «л», «т»).

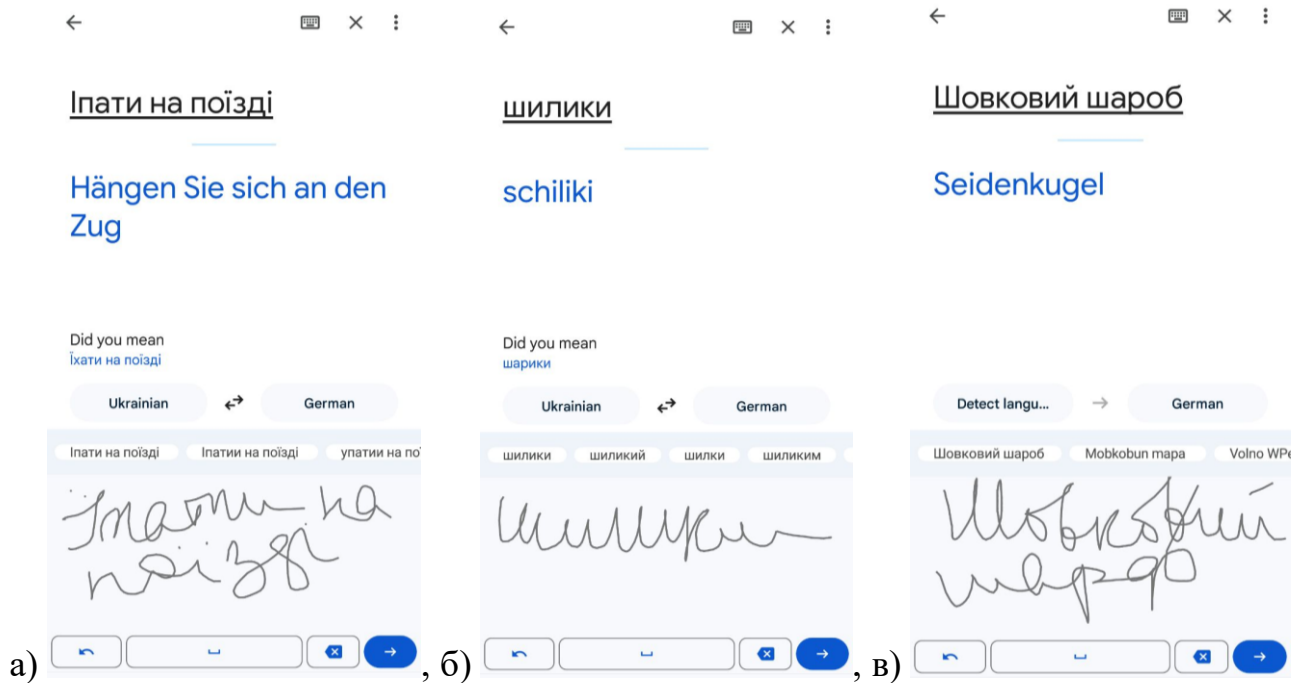


Рисунок 2.18. Результати експерименту з розпізнавання рукописного тексту за допомогою мобільного додатку Google (літери «ш», «ї», «р», «ф» та комбінації).

Фраза «Їхати на поїзді» не була розпізнана правильно (рисунок 2.18, а). Проблема спостерігалася з літерою «х», а літери «ї», «і» були розпізнані вірно. Як ми бачимо також на рисунку 2.18 (б, в) складнощі викликають літери «ш», «ф». Можна дійти висновку, що це пов'язано з наявністю однакових елементів у алгоритмах написання літер. Наприклад, літера «ф» у контексті алгоритму написання дійсно схожа на комбінацію «ар» (рисунок 2.19), особливо у стилі написання, що наближений до друкованої форми.

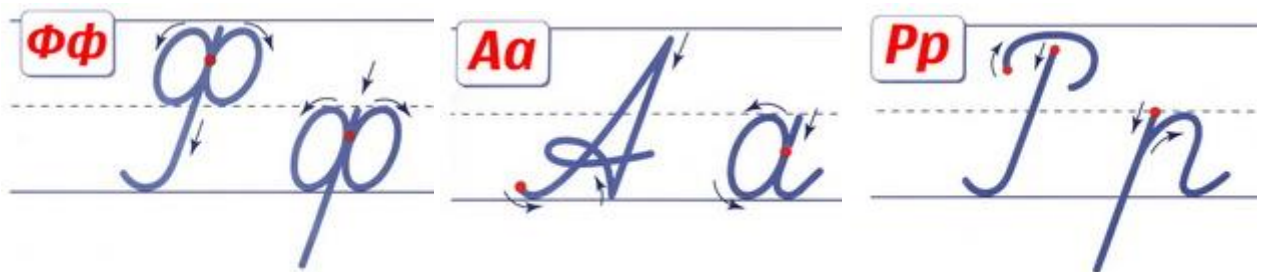


Рисунок 2.19. Алгоритм написання літер «ф», «а», «р».

Отже, завдання програми з розпізнавання рукописного тексту ускладнюється не тільки особливостями почерку, але й наявністю схожих елементів.

## ВИСНОВКИ

Проведене дослідження дозволило зробити наступні висновки:

1. Проблема розпізнавання рукописного тексту є актуальною, а рішення цієї проблеми є необхідним для багатьох галузь функціонування людини.
2. Розвиток технологій розпізнавання рукописного тексту відбувається у двох глобальних напрямках: офлайн, коли для розпізнавання надається готовий текст, та онлайн, який ще називають паралельним, коли процес розпізнавання тексту відбувається паралельно його створенню. Результати досліджень у вказаних напрямках значним чином відрізняються, що пов'язано, перш за все, з характерними особливостями рукописного тексту.
3. Незважаючи на те, що спроби з розпізнавання тексту беруть свій початок ще з 40х років, й кількість експериментів є значною, на даний момент не існує офлайн програмного комплексу, що б дозволяв отримати результати зі 100% ефективністю.
4. Варто відзначити, що у сфері онлайн розпізнавання рукописного тексту є значні результати. По-перше, технологія паралельного розпізнавання рукописного тексту широко застосовується у багатьох програмних комплексах й компонентах, що дозволяє вдосконалювати технологію й давати значний рівень ефективності. Онлайн розпізнавання рукописного тексту стало невід'ємною частиною комунікації людини з комп'ютерними технологіями.
5. Програмні компоненти компанії Google виявилися найбільш ефективними. При цьому, сервіси ідентифікують мову та регістр. Встановлені певні відмінності між браузерним та мобільним сервісом. Мобільний додаток не зміг розпізнати рукописний текст з фото, однак, введення тексту рукописним чином виявилось однаково ефективним як для мобільного, так й для сервісу у браузері. Спостерігалися складнощі з розпізнаванням літер «ш», «х» та «и», а літери, що

вважалися складними (і, д, б) були розпізнані коректно. Отже, складними є літери, що схожі за алгоритмом написання.

6. Проведений експеримент дозволив зробити висновок, що офлайн технології розпізнавання рукописного тексту ще недостатньо розвинуті, а онлайн технології дозволяють розпізнати рукописний текст з високим рівнем чіткості.

**ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ**

1. Богуславский А.А., Соколов СМ. Программно-аппаратные средства ввода зрительных данных в память персонального компьютера. Препринт ИПМ им. М.В.Келдыша РАН №52, Москва, 2002. - 22 с.
2. Болотова Ю.А., Спицын В.Г., Осина П.М. Обзор алгоритмов детектирования текстовых областей на изображениях и видеозаписях. Компьютерная оптика. 2017. – Т. 41, № 3. С. 441-452. DOI: 10.18287/2412-6179-2017-41-3-441-452.
3. Друки А.А., Милешин М.А. Алгоритмы распознавания рукописных подписей на основе нейронных сетей. Фундаментальные исследования. 2014. № 11-9. – С. 1906-1910. – URL: <http://www.fundamental-research.ru/ru/article/view?id=35866>.
4. Спицын В.Г., Болотова Ю.А., Фан Н.Х. Применение вейвлет-преобразования Хаара, метода главных компонент и нейронных сетей для оптического распознавания символов на изображениях в присутствии импульсного шума. URL: <https://cyberleninka.ru/article/n/raspoznavanie-simvolov-na-osnove-veyvlet-preobrazovaniya-metoda-glavnyh-komponent-i-neyronnyh-setey>.
5. Herbert F. The history of OCR, optical character recognition / Herbert F. Schntz – Manchester: Manchester Center, Vt., 1982 – 156 p.
6. Жихаревич С. Е., Остапов І. В. Аналіз методів розпізнавання символів тексту URL: <http://nti.khai.edu:57772/csp/nauchportal/Arhiv/REKS/2016/REKS516/Zhikharevich.pdf>.
7. Brakensiek, A. Off-line handwriting recognition using various hybrid modeling techniques and character N-grams [Text] / A. Brakensiek, J. Rottland, A. Kosmala,

- G. Rigoll // Proceedings of the Seventh International Workshop on Frontiers in Handwriting Recognition. — September 11-13. — 2000. — P. 343-352.
8. Jonsson, M. Ocr-algorithm for detection of subtitles in television and cinema [Text] / M. Jonsson, H. H. Bothe // Conference & Workshop on Assistive Technologies for People with Vision & Hearing Impairments Assistive Technology for All Ages CVHI. — 2007.
  9. Swype-клавіатура. [Електронний ресурс] – Режим доступу до ресурсу: <https://android.biblprog.org.ua/ua/swype-keyboard/>
  10. Cruz, R. M. O. CruzHandwritten Digit Recognition Using Multiple Feature Extraction Techniques and Classifier Ensemble [Text] / R. M. O. Cruz, G. D. C. Cavalcanti, T. I. Ren // 17th International Conference on Systems, Signals and Image Processing. — 2010. — P. 215-218.
  - 11.83 Lehal, G. S. Feature Extraction and Classification for OCR of Gurmukhi Script [Text] / G. S. Lehal, C. Singh // Vivek. — Vol. 12. — No. 2. — 1999. — P. 2-12.
  12. Liu, C.L. High Accuracy Handwritten Chinese Character Recognition Using Quadratic Classifiers with Discriminative Feature Extraction [Text] / Cheng-Lin Liu // Pattern Recognition, 2006. ICPR 2006. 18th International Conference on vol. 2. — 2006. — P. 942-945.
  13. Rico, J. R. Off-line cursive handwritten word recognition based on tree extraction and an optimized classification distance [Text] / J. R. Rico // National Symposium on Pattern Recognition and Image Analysis. — Vol. 2. — 1999. — P. 15-16.
  14. Steinert, R. Recognition of handwritten digits using sparse codes generated by local feature extraction methods [Text] / R. Steinert, M. Rehn, A. Lansner // ESANN'2006: 14th European Symposium on Artificial Neural Networks. — 2006. — P. 161-166.

15. Khushbu, M. Image Pre-processing on Character Recognition using Neural Networks [Text] / M. Khushbu, S. Mehta // International Journal of Computer Applications. — Volume 82. — No 13. — November 2013. — P. 11-15.
16. Перекладач від компанії Google. [Електронний ресурс] – Режим доступу до ресурсу:  
<https://translate.google.com.ua/?hl=ru&tab=rT&sl=uk&tl=uk&op=translate>
17. ABBYY FineReader [Електронний ресурс] – Режим доступу до ресурсу:  
<https://www.abbyy.com/en-ee/finereader/>
18. Convertio [Електронний ресурс] – Режим доступу до ресурсу:  
<https://convertio.co/ru/ocr/>
19. Img2Txt [Електронний ресурс] – Режим доступу до ресурсу:  
<https://img2txt.com/uk>
20. Методика навчання української мови з каліграфією: опорні конспекти лекцій (Частина 1). Методика навчання грамоти. Методика розвитку мовлення: для студентів денної та заочної форм навчання спеціальності 013 «Початкова освіта», освітній ступінь «бакалавр»/ укладач О.М.Хома.—Мукачево: МДУ, 2019.— 63 с.
21. Безкоровайна О. В. Прописи. Українська мова. Каліграфічний тренажер. – Х.: Навчальна література, 2016. – 64 с. – (Каліграфія).