

## **СТВОРЕННЯ ОФЛАЙН-ДОДАТКУ ДЛЯ ШВИДКОЇ ТРАНСКРИПЦІЇ ДРУКОВАНОГО ТЕКСТУ ТА ФОРМУЛЬНИХ ВИРАЗІВ**

Горгуль Д. О.

Науковий керівник – ст. викладач Гриньова О. Є.

Харківський національний університет радіоелектроніки, каф. ІІІ,  
м. Харків, Україна

тел. +38(050) 804-26-55

This paper proposes a solution that involves the use of optical character recognition (OCR) technology to convert printed text and certain kinds of formulas into digital text [1], which can then be edited and saved by the user. The application would be designed to work offline, without the need for an internet connection, and would prioritize speed and accuracy in the transcription process.

The proposed program would use a combination of machine learning algorithms, document writing rules, and natural language processing techniques to improve its accuracy and efficiency. Limiting the program to recognizing only printed text will increase the accuracy of printed text recognition.

В сучасному світі, де кількість друкованої інформації зростає з кожним днем, виникає потреба у швидкій транскрипції цієї інформації в електронний вигляд. Інформаційні технології розвиваються настільки швидко, що на сьогоднішній день можливо створити ефективний інструмент для автоматичної транскрипції друкованого тексту. Однак велика кількість існуючих додатків залежить від доступу до Інтернету, що ускладнює роботу з ними в офлайн-режимі, а також вимагають додаткової інформації про користувача та мають інші обмеження на їх використання.

Вирішення задачі з транскрипції в цілому поділяється на декілька етапів:

1. Класифікація символів.
2. Ідентифікація символічних регіонів.
3. Розширення символічних регіонів.
4. Послідовний збір текстових рядів.

Додаток пропонується використати як частину основного інструментарію операційних систем, на рівні з таким додатком як Microsoft Snipping Tool.

Розглянемо основні переваги запропонованого додатку:

– Програма працює без доступу до Інтернету, тому користувач може використовувати його навіть в умовах, коли немає Інтернету або при його низькій вартості.

– Програма забезпечує швидку транскрипцію друкованого тексту за допомогою технології OCR, яка працює локально. Немає необхідності робити API виклики, надсилаючи дані та чекаючи, поки модель надає резуль-

тат для відповіді. Це може бути критичною точкою для програм, які працюють у реальному часі.

- Для використання додатку потрібно лише настроїти комбінацію клавіш та виділити потрібну область для транскрипції.

- Технологія OCR дозволяє отримувати високу точність транскрипції друкованого тексту, що дозволяє значно зменшити кількість помилок.

Ця система також має декілька недоліків:

- Для підвищення точності та швидкості програми довелося урізати його функціонал до розпізнавання виключно друкованого тексту.

- При введенні правок до додатку потрібно проводити оновлення всіх екземплярів програми.

- Перші версії програми натреновані на невеликій збірці мов.

- Оскільки додаток працює не на дистанційному сервері, то він буде займати більше місця в порівнянні з онлайн рішеннями.

- Швидкість роботи додатку залежить від пристрою.

Останні два недоліки можна поліпшити за допомогою розвинутих алгоритмів оптимізації, та з використанням більш до апаратного рівня набору інструментів розробки.

Також треба розглянути конкурентні рішення:

- Mathpix Snipping Tool – найбільш схожий до нашого рішення інструмент. Він потребує реєстрації користувача, та дає всього десять спроб на добу. Перевагами додатку є те, ще він може читати друкований та рукописний текст, а також розрізняє різноманітні формульні вирази та, за потреби, пропонує їх рішення.

- ABBYY FineReader – дозволяє робити розпізнавання друкованого тексту з високою точністю, але може працювати лише на деяких форматах документів (частіше всього PDF).

- Google Cloud Vision API – є онлайн рішенням для транскрипції друкованого та рукописного тексту. Цей сервіс підтримує велику кількість мов, та гарантує точність транскрипції. Недоліком цього рішення є велика ціна використання, та питання конфіденційності інформації, що зберігається на серверах Google.

Список використаних джерел:

1. A Simple Equation Region Detector for Printed Document Images in Tesseract. Zongyi Liu, Ray Smith. 12<sup>th</sup> International Conference on Document Analysis and Recognition (ICDAR), 2013.