

УДК 004

Нестеров Г.Д.¹, Гофман Є.О.²

¹ студ. гр. КНТ-124м НУ «Запорізька політехніка»

² доц. НУ «Запорізька політехніка»

РОЗРОБКА ТА ДОСЛІДЖЕННЯ АЛГОРИТМІВ МАШИННОГО НАВЧАННЯ ДЛЯ РОЗПІЗНАВАННЯ РУКОПИСНОГО ТЕКСТУ

Об'єкт дослідження – процес розробки програмного забезпечення для розпізнавання рукописного тексту.

Предмет дослідження – програмні засоби для реалізації алгоритмів машинного навчання для розпізнавання рукописного тексту.

Мета роботи – створення програмного забезпечення для підвищення точності розпізнавання рукописного тексту.

Матеріали, методи та технічні засоби: html, css, мови програмування Python і JavaScript, TensorFlow, персональний комп'ютер під управлінням операційної системи Microsoft Windows 11.

У результаті виконання роботи проведено аналіз методів розпізнавання рукописного тексту та шляху їх розвитку, проведено детальне ознайомлення із предметною областю – розпізнаванням рукописного тексту, ознайомлення із найбільш важливими питаннями цієї проблемної області та проаналізовано існуючі рішення.

У роботі описано причину вибору технологій розробки, їх опис, переваги і недоліки.

У межах роботи мовою реалізації обрано Python, що зумовлено його широким застосуванням у галузі машинного навчання та наявністю потужного інструментарію для побудови і навчання нейронних мереж (TensorFlow, Keras, PyTorch). Для навчання нейронної мережі, було обрано платформу TensorFlow і, як середовище розробки, було використано Visual Studio Code.

У роботі проведено аналіз теоретичних основ машинного навчання та глибинних нейронних мереж, які застосовуються для задач оптичного розпізнавання тексту. Розглянуто сучасні архітектури моделей, зокрема рекурентні (RNN, LSTM) та згорткові нейронні мережі (CNN), а також їх комбінації, які забезпечують високу точність під час розпізнавання послідовностей символів. Також було описано процес збору і обробки вхідних даних, а також процес тренування нейронної мережі.

Окрім нейронної мережі, було спроектовано функціональну схему програми і Uml-діаграму композитної структури класів, які наглядно демонструють теоретичний та практичний потік роботи при виконанні завдання. Описано кожен використаний метод, кожен модифікацію і причини їх застосування.

Для підвищення зручності використання створено інтерактивний графічний інтерфейс, який дозволяє користувачу завантажувати зображення і переглядати результати розпізнавання з візуалізацією. Наведено потік функціонування програми і усі форми, які користувач може побачити.

Після проведення тестування, отримані результати підтверджують ефективність обраних методів і створюють основу для подальших досліджень у сфері інтелектуального аналізу зображень і документів. В порівнянні із аналогами, розроблена програма видає конкурентоспроможні результати. До того ж, розглянуто залежність результату роботи програми від пошкодженості вхідного зображення і втрати чіткості розпізнавання визначено як мінімальні і допустимі.

Актуальність та проблематика роботи полягає в зростаючому попиті на машинне навчання на ринку товарів ІТ галузі і, водночас, цифролізації документообігу і повсякденної передачі текстової інформації.

Наукова новизна одержаних результатів:

- удосконалено методику попередньої обробки рукописних зображень шляхом комбінування адаптивної бінаризації, нормалізації та контурного фільтрування, що забезпечує підвищення якості вхідних даних перед подачею в модель глибокого навчання;

- запропоновано модифіковану архітектуру рекурентної нейронної мережі, оптимізовану для розпізнавання дрібних рукописних символів шляхом поєднання згорткових блоків із шарами підвищеної чутливості до локальних особливостей зображення та використання довгої короткострокової пам'яті для збереження проміжних результатів розпізнавання;

- удосконалено підхід до навчання моделі за рахунок використання розширеного датасету для навчання, що підвищує стійкість мережі до різних типів рукописів;

- розроблено програмний застосунок із використанням власної моделі машинного навчання, що забезпечує інтеграцію процесів попередньої обробки, розпізнавання та візуалізації результатів.

В подальшому можна модернізувати програмний продукт в напрямку розширення функціоналу, наприклад, шляхом додавання різних мов для розпізнавання.

Галузь використання – програмний застосунок може бути застосований в багатьох галузях, зокрема, в організації навчального процесу закладів освіти, переведенні архівних документів державних підприємств в цифровий вигляд тощо.