

ТЕХНОЛОГІЯ ОПТИЧНОГО РОЗПІЗНАВАННЯ РУКОПИСНИХ СИМВОЛІВ В СИСТЕМАХ ЕЛЕКТРОННОГО ДОКУМЕНТООБІГУ

Вінницький національний технічний університет

Анотація

Запропоновано підхід по розпізнаванню рукописних текстових документів із використанням згорткової нейронної мережі та структурного методу.

Ключові слова: розпізнавання текстових символів, згорткова нейронна мережа, структурний метод.

Abstract

An approach to the recognition of handwritten text documents using a convolutional neural network and a structural method is proposed.

Keywords: text symbol recognition, convolutional neural network, structural method.

Вступ

Завдання обробки та розпізнавання зображення текстових документів є одним із актуальних напрямків інформаційних технологій. Це використовується в системах для введення тексту в електронному вигляді, розпізнавання різних бланків статистичного обліку, бланків податкових декларацій, банківських рахунків та інших типів документів [1, 2]. Методи автоматичного розпізнавання образів та їх реалізація у системах оптичного читання текстів (Optical Character Recognition, OCR) є однією із найбільш чудових технологій штучного інтелекту [3]. Завдання переведення інформації із друкованих паперових на електронні носії є дуже актуальним. Ще складнішим стає це завдання при виконанні автоматизованого розпізнавання рукописних текстів, яке в даний час поки ще не вирішено у повній мірі та є однією з найбільш актуальних та складних задач розпізнавання даних. Розгляду одного із підходів по виділенню та розпізнавання рукописних текстових документів присвячений даний матеріал.

Розпізнавання текстових символів

Теперішні відомі підходи для розпізнавання символів текстового документу є методи шаблонні, ознакові та структурні [4]. Також значних успіхів досягнуто із використанням підходів на основі нейронних мереж [5]. Кожен із цих підходів характеризується своєю складністю та ефективністю. Самим простим є шаблонний метод, у основі якого лежить послідовне порівняння отриманого зображення символу із еталонними зображеннями. Найкращі результати розпізнавання показують методи на основі використання нейронних мереж. Але при знаходженні у текстовому документі значної кількості рукописних символів ні один із цих методів не дає бажаного результату.

Для підвищення якості розпізнавання рукописних символів текстових документів пропонується об'єднати методи на основі нейронних мереж та структурного підходу. Пошук та розпізнавання символів у цифрових зображеннях виконується за ряд етапів. Технологічна послідовність розпізнавання рукописних символів текстових документів буде такою: попередня обробка зображення, бінаризація, сегментація, виділення каркасу символів та розпізнавання символів. Етап попередньої обробки є початковою фазою усіх алгоритмів розпізнавання символів. На цьому етапі покладається виконання таких завдань: підвищення якості зображення за рахунок фільтрації, видалення завад та інші операції, що мають на меті підвищити якість зображення.

Підготовка рукописного документа до розпізнавання включає виправлення, бінаризацію по порогу та видалення шумів. Порогова бінаризація — це процес відокремлення фону від об'єкта, у даному

випадку тексту. У результаті отримуємо чорний текст на білому тлі. Зниження шуму видаляє артефакти із зображення, не впливаючи на написаний текст. Проводиться сегментація рядків, слів, символів, щоб у кінцевому підсумку полегшити подальше розпізнавання символів. Елементарна сегментація слів працює за принципом, що інтервал між словами більший, ніж між буквами. Змінюючи кут нахилу, форму букв випрямляють для подальшого розпізнавання. Потім до тексту застосовуються фільтри, щоб видалити завади і визначити межі слів.

Виділення тексту на зображенні є одним із етапів розпізнавання. На цьому етапі роботи на попередньо бінаризованому зображенні виділяється безпосередньо область, на якій знаходиться текст, що підлягає розпізнаванню. Виконуються операції сегментації та нормалізації текстових фрагментів та аналіз текстового документа. Текст розділяється на зручні для аналізу складові частини. На даному етапі послідовно виконується поділ тексту на окремі рядки (сегментація рядків), потім поділ рядків на окремі слова (сегментація слів), а надалі поділ виділених слів на елементарні складові частини у вигляді символів. Потім здійснюється нормалізація символів до необхідних розмірів для виконання процесу розпізнавання. Для виконання ряду дій на цих етапах роботи даної послідовності були використані деякі вже розроблені програми із бібліотеки із відкритим кодом OpenCV [6].

На завершальному етапі виконуємо розпізнавання символів із використанням нейронної мережі глибокого навчання [5]. Попередньо відбувається її налаштування на розпізнавання рукописних текстових символів. Нерозпізнані символи надходять на інший модуль програми, де відбувається їх розпізнавання із використанням структурного підходу. У модулі додаткового розпізнавання попередньо не розпізнаних рукописних символів здійснюється їх розпізнавання із використанням структурного методу.

Важливою задачею для виділення та розпізнавання символів текстового зображення є розробка ефективного програмного забезпечення. Для вирішення цієї задачі створена програмна реалізація запропонованого підходу з використанням мови програмування Python [7] та модулів бібліотеки OpenCV, що дозволяють здійснити процес виділення та розпізнавання рукописних символів текстових документів.

Запропонований підхід може бути використаний у комп'ютерних системах розпізнавання рукописних символів за отриманим цифровим зображенням текстових документів.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Кристофер М. Бишоп. Распознавание образов и машинное обучение. М.: Диалектика, 2020. — 624 с.
2. Форсайт Д. Компьютерное зрение. Современный подход / Д. Форсайт, Ж. Понс. — М.: Вильямс, 2004. — 928 с.
3. Оптичне розпізнавання символів [Електронний ресурс] — Режим доступа: [https://ua.wikipedia.org/wiki/ Оптичне_розпізнавання_символів](https://ua.wikipedia.org/wiki/Оптичне_розпізнавання_символів).
4. Моругов А. М. Методы распознавания символов. / А.М. Моругов, С.В.Волков //Труды Международного симпозиума «Надежность и качество», 2017, том 1. С. 76—81.
5. Хайкин С. Б. Нейронные сети: полный курс. — 2-е изд. Пер. с англ. — М.: Издательский дом «Вильямс», 2006. — 1104 с.
6. OpenCV library [Электронный ресурс]. — Режим доступа: <https://opencv.org/>.
7. Коэльё Л. Построение систем машинного обучения на языке Python. / Коэльё Л. П., Ричерт В. — Перевод с английского. — М.: ДМК Пресс, 2015. — 364с.

Вадим Валерійович Осельський - студент групи ІКІ-20м факультету інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, Вінниця, e-mail: vadimoselskyi@gmail.com

Микола Андрійович Очкуров — старший викладач кафедри обчислювальної техніки, Вінницький національний технічний університет, м. Вінниця.

Vadym V. Oselskyi - students, Department of Information Technology and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: vadimoselskyi@gmail.com.

Mykola A. Ochkurov — Senior lecturer of the Computer Techniques Chair, Vinnytsia National Technical University, Vinnytsia.