

ВИКОРИСТАННЯ НЕЙРОННОЇ МЕРЕЖІ ДЛЯ РОЗПІЗНАВАННЯ РУКОПИСНИХ ЦИФР

Вінницький національний технічний університет

Анотація

Запропоновано підхід по розпізнаванню рукописних цифр із використанням згорткової нейронної мережі.

Ключові слова: розпізнавання символів, рукописні цифри, згорткова нейронна мережа.

Abstract

An approach to recognizing handwritten digits using a convolutional neural network is proposed.

Keywords: character recognition, handwritten digits, convolutional neural network.

Вступ

Завдання переведення інформації із паперових документів на електронні носії є актуальним у рамках потреб, що виникають в різних системах електронного документообігу. Сучасні інформаційні технології суттєво спрощують доступ до різноманітних інформаційних ресурсів за виконання такої умови, що вони будуть переведені в більш зручний для оброблення електронний вигляд. Найбільш швидким й простим є здійснення операції сканування текстових документів із використанням сканерів. У підсумку отримуємо цифрове зображення початкового документа в вигляді нового графічного файлу. Але значно кращим, в порівнянні з його графічним представленням, є текстове подання цієї інформації. Цей варіант дозволяє досить суттєво зменшити витрати на обсяги зберігання й передачі інформації, та створює додаткові можливості для реалізації усіх можливих сценаріїв використання й аналізу електронних документів. Тому із практичної точки зору значний інтерес представляє саме переведення наявних паперових носіїв у текстові електронні документи [1].

Досить складним стає це завдання при обробці документів шляхом автоматизованого розпізнавання рукописних текстів у тому числі і рукописних цифр, яке в даний час поки ще так і не вирішено в повному обсязі та вважається одним із найбільш актуальних й складних завдань розпізнавання даних [2]. Розгляду одного із підходів по виділенню та розпізнаванню рукописних цифр текстових документів присвячений даний матеріал.

Розпізнавання рукописних цифр текстових документів

Для створення програмного засобу по розпізнаванню текстових документів було вирішено використати згорткову нейронну мережу [3]. Структура програми для розпізнавання рукописних цифр складається із трьох базових модулів й одного допоміжного модуля. Ними є модуль попередньої обробки текстового зображень, модуль сегментації текстового документа та модуля розпізнавання рукописних цифр на основі нейронної мережі. У складі програми є ще один допоміжний модуль для початкового налаштування нейронної мережі для здійснення процесу розпізнавання рукописних цифр.

Модуль попередньої обробки текстового зображення виконує такі дії. Отримуємо зображення рукописного тексту, здійснюємо фільтрацію від можливих завад та морфологічних дефектів, переводимо зображення тексту з отриманого кольору у бітовий формат та почергово виділяємо області, що відповідають окремим сторінкам текстового документа. Надалі розділяємо виділену сторінку тексту на окремі рядки із використанням гістограм вертикального розподілу яскравості зображення текстового документа. Наступним кроком є розподіл рядків на окремі слова чи групи цифр із використанням гістограми розподілу яскравості виділеного рядка тексту по горизонталі. Надалі розподіляємо виділені групи цифр на окремі рукописні символи та визначаємо їх положення. Після

розподілу груп цифр на окремі символи із цифр виконуємо їх нормалізацію до необхідного розміру для подальшого виконання процесу розпізнавання рукописних цифр на основі згорткової нейронної мережі. Процес розпізнавання рукописних цифр відбувається із використанням нейронної мережі глибокого навчання.

Для налаштування модуля підготовки нейронної мережі для подальшої роботи використовуємо базу даних еталонних символів MNIST, до складу якої входить набір із базових символів цифр із значним розмаїттям різного їх написання, щоб настроїти згорткову нейронну мережу таким чином, щоб вона здійснювала процес розпізнавання рукописних цифр [4].

Наступним кроком обробки текстового документа буде збереження кожного з виділених рукописних чисел у масштабі 28x28 біт, оскільки база даних довідкових чисел MNIST працює в такому масштабі, як і нейронна мережа для розпізнавання рукописних цифр.

Також окрім обробки вхідних даних необхідно натренувати модель нейронної мережі. Для цього створюється великий набір даних, що складається з різних варіантів написання цифр. Для правильного використання моделі архітектури Faster R-CNN спочатку тренується базова нейронна мережа, що слугує класифікатором. Згодом тренується основна архітектура з натренованою базовою моделлю. Вихідний результат формується з двох складових частин: обмежувальної рамки об'єкту, тобто чотирикутника який обмежує об'єкт так, що жоден піксель не виходить за координати рамки; класу об'єкту, тобто його назви символу присутньому на зображенні.

Тренування нейронної мережі забезпечується завдяки модулю train, що приймає як аргумент зкомпільований об'єкт моделі. На початку функції у консоль виводяться обрані гіперпараметри. Ініціюється два об'єкти датасетів, перший відповідає за тренувальний набір даних, другий — за валідаційний, що є меншим за розміром та містить зразки, жодний з яких не є присутнім у тренувальному наборі, він необхідний для перевірки точності моделі на даних, що відрізняються від тренувальних, а отже визначають реальну приблизну точність в умовах використання вже натренованої моделі. Використання валідації може також допомогти при відслідковуванні проблеми перенавчання, тобто високої точності на відомих даних та низької при використанні даних, що невідомі моделі.

Перевірка якості розпізнавання тестування проводилась у два етапи. На першому етапі було перевірено точність відображення груп чисел у числовому форматі, а також перевірялися обмежувальні рамки розпізнаних рукописних чисел за допомогою модуля ImageDraw бібліотеки Pillow. Наступним протестованим модулем був структурний модуль, який використовує сформовані операції для представлення окремих цифр у документі.

Висновки

Розроблений програмний продукт може використовуватися в засобах розпізнавання рукописних цифр у текстових документах.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Жихаревич В. В. Аналіз методів розпізнавання символів тексту / В. В. Жихаревич, С. Е. Остапов, І. В. Миронів // *Радіоелектронні і комп'ютерні системи*. 2016, № 5. — С. 137 — 142.
2. Коротун О. В. Система розпізнавання рукописних цифр з оцінкою якості. / О. В. Коротун, Г. В. Марчук, Д. К. Марчук, О. В. Талавер // *Технічна інженерія*. № 1 (85) 2020. — с. 135— 146.
3. Субботін С. О. Нейронні мережі: теорія та практика: навч. посіб. / С. О. Субботін. — Житомир : Вид. О. О. Євенок, 2020. — 184 с.
4. LeCun, Yann; Corinna Cortes; Christopher J.C. Burges. MNIST handwritten digit database. <http://yann.lecun.com/exdb/mnist/>.

Безруков Владислав Олександрович — студент групи 2КІ-236, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, Вінниця, e-mail: bezrukovvladyslav8578212@gmail.com

Очкуров Микола Андрійович — старший викладач кафедри обчислювальної техніки, Вінницький національний технічний університет, Вінниця, e-mail: ochkurovma.50@ukr.net

Bezrukov Vladyslav O. — students, Department of Information Technology and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: bezrukovvladyslav8578212@gmail.com

Ochkurov Mykola A. — Senior lecturer of the Computer Techniques Chair, Vinnytsia National Technical University, Vinnytsia, e-mail: ochkurovma.50@ukr.net