

РОЗПІЗНАВАННЯ ТЕКСТУ НА ЗОБРАЖЕННЯХ ЗА ДОПОМОГОЮ EASYOCR

Вінницький національний технічний університет

Анотація

Проведено аналіз EasyOCR - відкритої бібліотеки з оптичного розпізнавання символів на основі глибокого навчання. Описуються основні принципи роботи EasyOCR. Також розглядаються сфери застосування EasyOCR та наведено приклади використання.

Ключові слова: EasyOCR, оптичне розпізнавання символів, глибоке навчання

Abstract

The analysis of EasyOCR - an open library for optical character recognition based on deep learning - was carried out. The basic principles of EasyOCR operation are described. Areas of application of EasyOCR are also considered and examples of use are given

Keywords: EasyOCR, optical character recognition, deep learning

Вступ

Оптичне розпізнавання символів (OCR) є однією з ключових технологій для автоматичного розпізнавання та перетворення тексту зі сканованих документів, фотографій та інших зображень на електронний формат. В сучасному світі, де інформація стає все більш доступною та широко розповсюдженою, OCR є незамінним інструментом для забезпечення швидкого та ефективного перетворення документів в цифровий формат. EasyOCR - це одна з найбільш ефективних та простих у використанні бібліотек OCR, яка базується на сучасних методах машинного навчання та забезпечує високу точність розпізнавання тексту на зображеннях. Окрім того, EasyOCR підтримує розпізнавання багатомовного тексту, що робить його особливо корисним для роботи з документами у різних мовах.

Огляд моделей для навчання EasyOCR

Convolutional Neural Networks (CNN): EasyOCR використовує навчальні моделі, засновані на згорткових нейронних мережах, для розпізнавання тексту на зображеннях. Ці моделі можуть навчатися на великій кількості зображень, що дозволяє досягнути високої точності розпізнавання тексту.

Recurrent Neural Networks (RNN): EasyOCR використовує також навчальні моделі на основі рекурентних нейронних мереж для розпізнавання тексту, зокрема Long Short-Term Memory (LSTM) та Gated Recurrent Units (GRU). Ці моделі добре працюють з послідовністю символів, що дозволяє їм досягнути високої точності розпізнавання тексту з більш складними мовними конструкціями.

Connectionist Temporal Classification (CTC): EasyOCR використовує алгоритм CTC для розпізнавання тексту. Цей алгоритм дозволяє розпізнавати текст на зображенні без необхідності визначення де точно починається та закінчується кожен символ.

Adaptive Thresholding: EasyOCR використовує адаптивне порогове значення для підвищення якості зображення та покращення точності розпізнавання тексту.

```
import easyocr
import cv2
from matplotlib import pyplot as plt
import numpy as np
```

```
IMAGE_PATH = 'sigh.jpeg'
```

```
reader = easyocr.Reader(['en'])
result = reader.readtext(IMAGE_PATH)
result
```

Рис. 1 - Фрагмент коду для обробки зображення мовою програмування Python

EasyOCR використовує нейронну мережу для розпізнавання тексту на зображенні. Кроки, які виконуються в цій функції, можна описати наступним чином:

- Зображення завантажується та перетворюється в чорно-біле, якщо воно не чорно-біле за замовчуванням.
- Зображення обробляється з використанням різноманітних методів, таких як згладжування, бінаризація та знешумлення, щоб покращити розпізнавання тексту на зображенні.
- Зображення поділяється на регіони, які містять текст.
- Кожен регіон тексту передається нейронній мережі, яка використовує навчальні дані для розпізнавання символів на зображенні.
- Найбільш ймовірний результат розпізнавання для кожного регіону тексту повертається в вигляді рядка символів, разом з координатами регіону на зображенні.

Кожен крок цього процесу може бути налаштований з допомогою відповідних параметрів, щоб підібрати оптимальні настройки для конкретного випадку.

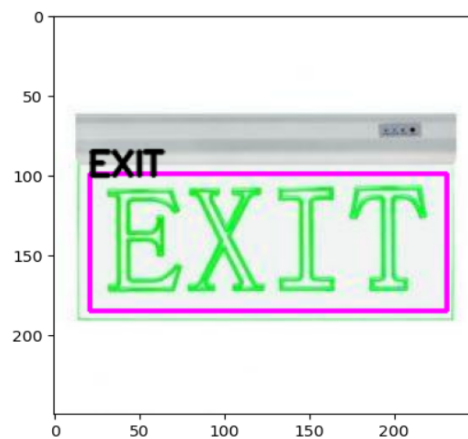


Рис. 2 – Приклад розпізнавання тексту за допомогою EasyOCR

Висновки

Системи OCR, мають великий потенціал застосування в різних галузях, що вимагають обробки великої кількості текстової інформації. Наприклад, в фінансовій сфері OCR може допомогти автоматизувати обробку банківських документів, в медицині - аналізувати медичні документи та рецепти, а в юридичній галузі - швидко та точно перетворювати документи на електронний формат. У зв'язку з тим, що EasyOCR є безкоштовним та має відкритий код, він є доступним для використання в різних проектах та дослідженнях. Системи OCR, зокрема

EasyOCR, становлять важливий інструмент для автоматизації процесів обробки текстової інформації та покращення продуктивності в різних галузях діяльності.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. SZELISKI, Richard. Computer vision: algorithms and applications. Springer Nature, 2022.
2. Mori, Shunji, Ching Y. Suen, and Kazuhiko Yamamoto. "Historical review of OCR research and development." Proceedings of the IEEE 80.7 (1992): 1029-1058.
3. Mithe, Ravina, Supriya Indalkar, and Nilam Divekar. "Optical character recognition." International journal of recent technology and engineering (IJRTE) 2.1 (2013): 72-75.

Водолазська Дар'я Вікторівна – студентка групи ІСП-19б, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, Вінниця, e-mail: dvodolazska@gmail.com.

Науковий керівник: **Крупельницький Леонід Віталійович** – к.т.н., доцент кафедри обчислювальної техніки, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, м.Вінниця

Vodolazska Daria V. – student of group 1SP-19b, Faculty of Information Technologies and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: dvodolazska@gmail.com.

Scientific adviser: **Krupelnitskyi Leonid V.** – and. Sc. (Eng.), Assistant Professor of the Computer Techniques Chair, Vinnytsia National Technical University, Vinnytsia