

МЕТОДИ ОПТИЧНОГО РОЗПІЗНАВАННЯ СИМВОЛІВ

Хоча більшість документів на сьогоднішній день складається на комп'ютері, завдання створення електронного документообігу ще далеко не повністю реалізовано. Як правило, наявні системи охоплюють діяльність однієї організації, а обмін даними між організаціями здійснюється за допомогою традиційних паперових документів.

Зі зростанням присутності цифрових носіїв у 21 столітті відбулося притаманне зростання потреби в оцифрованих документах. Документи, що зберігаються в цифровому вигляді, мають величезні переваги перед своїми аналогами в «реальному світі», зокрема щодо фізичного простору, який вони займають, та безпеки, пов'язаної з їх використанням. Як результат, аналіз документів за допомогою штучного інтелекту для оцифрування документів є невід'ємною частиною комп'ютерного зору і став сферою досліджень, що швидко розвиваються.

Оптичне розпізнавання символів, або OCR, є ключовим елементом оцифрування документів. Виявлення тексту із зображень документів дає змогу алгоритмам обробки природної мови розшифрувати текст і зрозуміти, що передає документ. Крім того, текст можна легко перекладати кількома мовами, що робить його легко інтерпретованим будь-кому. Проте OCR не обмежується виявленням тексту лише із зображень документів. Нові алгоритми OCR використовують комп'ютерний зір і NLP (обробка природної мови) для розпізнавання тексту з назв продуктів супермаркетів, дорожніх знаків і навіть з рекламних щитів, що робить їх ефективним перекладачем.

Сучасні методи OCR використовують алгоритми визначення тексту як відправну точку.

Найсучасніші нейронні мережі стали надзвичайно вдалимими для визначення тексту в документах і зображеннях, навіть якщо він нахилений, повернутий або перекошений.

Хоча традиційні підходи, засновані на машинному навчанні, швидко розвиваються, вони займають значно більше часу на виконання і легко випереджають алгоритми глибокого навчання як за точністю, так і за швидкістю висновку. Традиційні підходи до розпізнавання тексту проходять через серію етапів попередньої обробки, де перевіряється документ очищається від шуму і приводиться до вдалого виду. Після цього документ двійковується для подальшого виявлення контурів, щоб допомогти у виявленні рядків і стовпців. Нарешті, символи, що будують лінії, витягуються, сегментуються та ідентифікуються за допомогою різних алгоритмів машинного навчання, таких як найближчі сусіди та машини опорних векторів.

Хоча алгоритми вони чудово працюють із простими наборами даних OCR, такими як легко розпізнані друковані дані та рукописні дані, вони упускають багато функцій, через що вони не працюють під час роботи зі складними наборами даних.

Методи глибокого навчання можуть ефективно витягувати велику кількість функцій, що робить їх кращими за аналоги машинного навчання. Алгоритми, які поєднують підходи на основі NLP, були особливо успішними в забезпеченні чудових результатів для розпізнавання та виявлення тексту в дикій природі. Крім того, ці методи забезпечують наскрізний конвеєр виявлення, який звільняє їх від довгих кроків попередньої обробки.

Як правило, методи OCR включають підходи, засновані на баченні, що використовуються для виділення текстових областей і передбачення координат рамки для них.

Дані обмежувальної рамки та елементи зображення потім передаються в алгоритми обробки мови, які використовують RNN, LSTM і трансформатори для декодування інформації на основі функцій у текстові дані. Сьогодні оптичне розпізнавання символів відіграє ключову роль у процесі цифрової трансформації для багатьох компаній, допомагаючи їм безпечно зберігати свої дані та легше отримувати інформацію. Маркетингові агентства також використовують алгоритми OCR, щоб збільшити залучення клієнтів і збільшити продажі, забезпечуючи безперебійну роботу покупців.

Окрім користі для бізнесу, OCR допомагає доквіллю, зменшуючи кількість повторюваних друкованих копій важливих документів, заощаджуючи таким чином папір.

І останнє, але не менш важливе, OCR допомагає перекладати письмовий текст багатьма мовами, що збільшує доступність документів і допомагає подолати мовний розрив.

Література

1. Обробка природної мови [Електронний ресурс] – Режим доступу до ресурсу: <https://uk.wikipedia.org/wiki/NLP>
2. Квасников В.П. Улучшение визуального качества цифрового изображения путем поэлементного преобразования / Квасников В.П., Дзюбаненко А.В., 2009.