

РОЗВИТОК ВИКОРИСТАННЯ НЕЙРОННИХ МЕРЕЖ ОПТИЧНОГО РОЗПІЗНАВАННЯ ДЛЯ ОБРОБКИ ІНФОРМАЦІЇ

Єна О.О.

*Національний технічний університет
«Харківський політехнічний інститут», м.Харків*

У роботі розглянуто розвиток використання нейронних мереж задля оптичного розпізнавання з метою вилучення і обробки інформації.

Обсяги інформації в наш час неперервно збільшуються, що призводить до зростання значущості технологій обробки та видобутку тексту з різних джерел. Оптичне розпізнавання символів широко використовується в сферах медицини, права, освіти та багатьох інших.

Найпоширеніші системи в оптичному розпізнаванні в даний момент ґрунтуються на використанні глибоких нейронних мереж. Більшість рішень на ринку використовують поєднання конволюційних та рекурентних нейронних мереж, а також інколи нейронні мережі довгої короткочасної пам'яті.

Проте, ще в 2017 році була представлена нова архітектура трансформер моделей в статті «Attention Is All You Need» від Google. Як зазначають автори, «Для завдань перекладу трансформер можна навчити значно швидше, ніж на основі архітектур на рекурентних або згорткових шарах.»[1]. Згодом вийшла стаття де була представлена ефективна модель оптичного розпізнавання побудована на трансформер архітектурі – TrOCR. В статті чітко вказується: «Експерименти показують, що модель TrOCR перевершує актуальні найсучасніші моделі на друкованому, рукописному і завдання на розпізнавання тексту сцени» [2]. Зараз вже можна чітко спостерігати зацікавленість в побудуванні продуктів на новій архітектурі.

Найвідоміші рішення на ринку наразі використовують попередні до трансформер моделей архітектури. Постає проблема, коли попит на нові технології не зустрічає доступний звичайним користувачам відповіді. Крім того, наприклад, відома програма FineReader від російської компанії АBBY, яка пропонує оптичне розпізнавання, попри заборону досі розповсюджується в Україні[3], що має викликати стурбованість, а також має бути мотивацією для розробки власних продуктів на новітніх технологіях.

Отже, попит на програми оптичного розпізнавання символів тільки зростає і розробка продуктів на відносно новій архітектурі трансформер моделей має стати вкрай важливим і плідним напрямком роботи фахівців.

Література:

1. Attention Is All You Need [Електронний ресурс] / [A. Vaswani, N. Shazeer, N. Parmar та ін.] // 31st Conference on Neural Information Processing Systems. – 2017. – Режим доступу до ресурсу: <https://arxiv.org/pdf/1706.03762.pdf>.
2. TrOCR: Transformer-based Optical Character Recognition with Pre-trained Models [Електронний ресурс] / [M. Li, T. Lv, J. Chen та ін.]. – 2022. – Режим доступу до ресурсу: <https://arxiv.org/pdf/2109.10282.pdf>.
3. Карпенко О. Російська АBBY досі працює в Україні, а держоргани закуповують її софт [Електронний ресурс] / Ольга Карпенко // ain. – 2022. – Режим доступу до ресурсу: <https://ain.ua/2022/08/08/rosijska-abbyy-dosi-praczuje-v-ukrayini/>.