

ПРОГРАМНІ ЗАСОБИ ДЛЯ РОЗПІЗНАВАННЯ МАТЕМАТИЧНИХ ФОРМУЛ

Овсяніков В.В., Татарінова О.А.

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

Доповідь присвячено опису розробленого авторами програмного засобу для реалізації комп'ютерного зору, що застосовується для розпізнавання, як і окремо друкованих символів, так і цілих текстів, що можуть містити математичні формули та подальшого збереження результуючого документа у форматі Тех. Наразі існує велика кількість наукових документів, що не мають електронної копії в цілому, або ж не мають копії у загально прийнятому форматі написання наукових текстів. Оскільки мануальне переписування такої кількості файлів майже неможливе, або ж може бути досягнуте із застосуванням значної кількості людських ресурсів, розробка програмного засобу для автоматизації цієї задачі для підтримання бази наукових документів у відповідному вигляді стає надзвичайно важливим та актуальним завданням.

У даному програмному забезпеченні реалізовано можливість розпізнавання друкованих символів латиниці, кирилиці, літер грецького алфавіту та спеціальних математичних символів. Для цього застосовуються багат шарова згортальна нейронна мережа, побудована за допомогою бібліотеки машинного навчання Кєрас, та додаткові валідаційні євристики. Для підвищення якості розпізнавання нейронної мережі застосовано складний механізм преобробки зображень, що допомагає видалити з них шуми, виключити похибки, пов'язані з нахилом символів, коректувати дефекти символів, обумовлені якістю вхідного зображення. Також реалізовано механізми збирання окремих символів в слова або ж математичні формули, відтворення положення знаків індексів та ступенів, формування звичайних дробів та виразів під знаком кореня.

Для демонстрації можливостей розробленого програмного забезпечення додано графічний інтерфейс користувача, за допомогою якого можливо ще до початку розпізнавання обрати та оглянути вхідне зображення. Під час тестування програмного засобу було проведено розпізнавання зображень різних типів: повністю текстуальні, математичні формули без тексту, математичні формули, які знаходяться між блоками тексту. На даний момент алгоритм працює лише з зображеннями, збереженими у високій якості та планується вдосконалення алгоритму для розпізнавання зображень, збережених в низькій якості.