

Конфигурации ScanSSD и BTTR для распознавания математических документов в чат-ботах для обучения студентов бакалавриата математике

Скитская А.В (ИТМО)

Научный руководитель – кандидат философских наук, доцент Блейхер О.В (ИТМО)

Введение. Современные автоматизированные системы обучения математике в университетах становятся всё более распространёнными и активно развиваются. При создании таких систем разработчики сталкиваются с задачей качественного распознавания математических документов, для решения которой ключевую роль играет компьютерное зрение (КЗ). Распознавание математических документов является одной из сложных задач в области компьютерного зрения, так как такие документы содержат структурированную информацию, включая текст, формулы, графики и таблицы. Для успешного решения данной задачи необходима разработка и применение специализированных моделей КЗ. В данной работе была рассмотрена возможность распознавания математических документов на основе двух моделей компьютерного зрения - Scanning Single Shot Detector [1] (ScanSSD) и Bidirectional Trained Transformer [2] (BTTR) в контексте выбора архитектур моделей и способов их взаимодействия для достижения высокоточного результата.

Основная часть. В данной работе предлагается рассмотреть интеграцию технологий ScanSSD и BTTR в образовательные чат-боты для обучения математике. ScanSSD, являющийся моделью для обнаружения объектов, позволяет эффективно идентифицировать математические символы и формулы из изображений. В то же время, BTTR может обрабатывать текстовые элементы, позволяя преобразовывать распознанный текст в данные, легко понимаемые как ботами, так и пользователями.

Предлагаемое решение основывается на создании гибридной последовательной модели. В начале предполагается использовать модель ScanSSD для обнаружения математических формул в документах. После успешного обнаружения будет использована модель BTTR для дальнейшего распознавания и преобразования данных в формат LATEX.

Внедрение системы предполагается для проверки работ студентов в трех университетах - ИТМО, РЭУ, РАУ.

Выводы. Разработана гибридная модель на основе конфигураций ScanSSD и BTTR.

Список использованных источников:

1. Mali, Parag, et al. "ScanSSD: Scanning Single Shot Detector for Mathematical Formulas in PDF Document Images." ArXiv:2003.08005 [Cs], Mar. 2020
2. Wenqi Zhao, Liangcai Gao, Zuoyu Yan, Shuai Peng, Lin Du, Ziyin Zhang, Handwritten Mathematical Expression Recognition with Bidirectionally Trained Transformer, Document Analysis and Recognition – ICDAR 2021: 16th International Conference, Lausanne, Switzerland, September 5–10, 2021, Proceedings, Part II, pp. 570-584