

ИСПОЛЬЗОВАНИЕ YANDEX VISION OCR И ML KIT С OPENCV ДЛЯ РАСПОЗНАНИЯ ТЕКСТА

Оптическое распознавание текста (OCR) – важная задача для мобильных приложений, позволяющая автоматизировать ввод данных с фотографий. Особый интерес представляет распознавание кириллического текста. В работе рассматривается комбинированный подход к OCR, объединяющий Yandex Vision, Google ML Kit и OpenCV.

Yandex Vision OCR – облачный API, обеспечивающий высокоточное распознавание текста на более чем 48 языках. Для его использования необходимо интернет-соединение и API-ключ.

Google ML Kit предоставляет возможность локального распознавания текста, что обеспечивает быструю работу и повышенную конфиденциальность.

OpenCV применяется для предобработки изображений: удаления шумов, повышения контрастности и преобразования в градации серого.

Изображение, полученное с камеры или галереи, обрабатывается с использованием, затем передаётся в один из OCR-модулей. При наличии соединения с интернетом выполняется запрос к Yandex Vision OCR. В оффлайн-режиме используется Google ML Kit.

В дипломном проекте BeaX реализована именно такая система. Приложение выполняет предварительную обработку фотографий косметических продуктов с помощью OpenCV, далее осуществляется распознавание текста средствами Yandex Vision или ML Kit.

Предобработка изображений оказалась ключевым этапом: без неё точность заметно снижалась. Yandex Vision превосходит по качеству локальные решения, а ML Kit обеспечивает автономность. Подход, реализованный в проекте BeaX, позволяет выполнять сканирование описания косметического продукта и разбор его состава.

ЛИТЕРАТУРА

1. Yandex Vision OCR [Электронный ресурс] – Режим доступа: <https://cloud.yandex.com/en-ru/services/vision> (дата обращения 12.04.2025).
2. Google ML Kit [Электронный ресурс] – Режим доступа: <https://developers.google.com/ml-kit> (дата обращения 12.04.2025).