

# ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ОПТИЧЕСКОГО РАСПОЗНАВАНИЯ СИМВОЛОВ ДЛЯ МОБИЛЬНЫХ УСТРОЙСТВ С ПРИМЕНЕНИЕМ ОБЛАЧНЫХ ВЫЧИСЛЕНИЙ

*Д.Р. Мухтар, магистрант гр. 8ИМ02  
Томский политехнический университет  
E-mail: drm7@tpu.ru*

## **Введение**

В современном мире широко распространено хранение и обмен текстовой информацией посредством мобильных устройств в виде растровых графических файлов типа JPEG, JPG и PNG. Недостатком использования графических файлов является отсутствие возможности редактирования и копирования текста для использования в других приложениях. Для получения текстовой информации из графических файлов существует множество видов программ оптического распознавания символов, однако на мобильном устройстве не всегда имеется память для их установки и использования [1].

Поэтому целью работы являлось написание программного обеспечения, которое предоставило бы возможность распознать текст из графических файлов через мессенджер Telegram.

## **Описание алгоритма**

Для решения задачи оптического распознавания символов было разработано консольное приложение, работающее на базе программной платформы .NET 6, в среде разработки Microsoft Visual Studio 2022 Professional на объектно-ориентированном языке программирования C#. Подключение и работа с сервисом Telegram осуществляется через библиотеку C# Telegram. Обработка и распознавание символов проводится с использованием технологии компьютерного зрения Microsoft Azure, обеспечивающей доступ к расширенным алгоритмам обработки изображений.

Для создания программы были получены токен и ключ подписки для Telegram. Сперва написанный бот подключается к сервису Telegram используя эти данные аутентификации. Затем с применением технологии длинных опросов, программа ожидает получения данных от сервиса Telegram. Сервис Telegram получает запрос от программного обеспечения и отправляет ответ, когда произойдет какое-либо событие, а именно будет получен файл от пользователя. Данный файл сохраняется в базе данных Telegram и к нему присваивается уникальный ключ. Затем сервис Telegram отправляет данные о файле боту, который отвечает новым запросом на получение строки с URL-адресом файла хранящимся в базе Telegram. После получения строки с URL-адресом файла, отправляется запрос содержащий данную строку на сервис Azure, в котором происходит анализ и обработка файла.

Существует множество алгоритмов оптического распознавания символов. Все алгоритмы состоят из блоков предобработки изображения, выделения признаков, распознавания символов и постобработки результатов распознавания. Сервис Azure использует внутренний алгоритм, разработанный компанией Microsoft с применением ИИ и нейронных сетей.

После того как облачный сервис Azure закончит обработку изображения, он отправляет текст в виде строк обратно боту, который в свою очередь пересылает эти данные сервису Telegram для отображения пользователю.

Пример отправки изображения показан на рисунке 1.

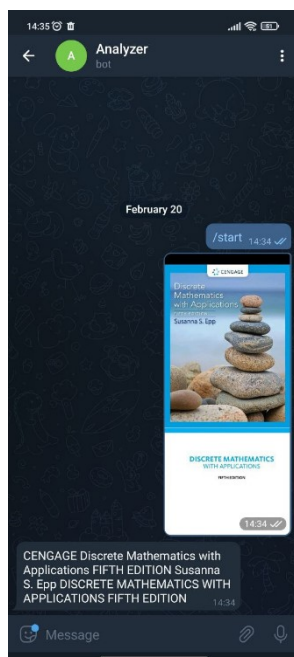


Рис. 1. Отправка изображения и получение ответа от сервиса.

Чтобы отправить файл с текстовой информацией, необходимо выбрать снимок экрана или фотографию текста и отправить его боту в мессенджере Telegram.

### Тестирование алгоритма поиска

Чтобы оценить производительность работы бота, были отправлены снимки экрана имеющие текстовую информацию с различным количеством символов. Тестирование производилось с устройства с разрешением экрана в 1080 x 2400 пикселей и при отправке изображения с количеством символов выше 4000, облачный сервис не производил обработку, что говорит о лимите на количество символов. Результаты представлены в таблице 1.

Таблица 1. Показатели производительности при различном количестве символов n

№ символов на снимке	Время выполнения распознавания (миллисекунд)
n=129	2850.7 мс.
n=768	3107.9 мс.
n=1748	2985.9 мс.
n=3046	3357.0 мс.
n=3644	2861.6 мс.

### Заключение

В результате проведения тестирования можно сделать вывод о том, что разработанная система OCR является работоспособной, а скорость обработки колеблется в районе трёх миллисекунд. Тем не менее система не функционирует, когда количество символов в изображении выше 4000.

### Список использованных источников

4. Михеева Е.В. Информационные технологии в профессиональной деятельности.– Проспект 2013.– 419 с.
5. Оптическое распознавание текста. [Электронный ресурс]. – URL: <https://docs.microsoft.com/ru-ru/azure/cognitive-services/computer-vision/overview-ocr> (дата обращения 20.02.2022).
6. Приложения реального времени и Polling ,Long Polling , WebSockets, Server-Sent Events SSE, WebRTC [Электронный ресурс]. – URL: <https://intellect.icu/prilozheniya-realnogo-vremeni-i-polling-long-polling-websockets-server-sent-events-sse-webrtc-7016> (дата обращения 20.02.2022).
7. Искусственный интеллект .NET. Речь, язык и поиск .– ДМК-Пресс, 2018 г. .– 298 с