

УДК 004.93

СРАВНИТЕЛЬНЫЙ АНАЛИЗ OCR-РАСШИРЕНИЙ ДЛЯ БРАУЗЕРОВ СЕМЕЙСТВА CHROMIUM

Ким Д.В.

Магистрант,

ФГБОУ ВО «Калужский государственный университет им К.Э. Циолковского»

Калуга, Россия

Кузнецова В.И.

к.пед.н., доцент

ФГБОУ ВО «Калужский государственный университет им К.Э. Циолковского»

Калуга, Россия

Аннотация

Статья описывает проблему извлечения текста с изображений и видео и направлена на исследование рынка OCR-расширений для браузеров. Проводится анализ существующих предложений, выделяются критерии для сравнения: способ обработки данных, дополнительные функции, конфиденциальность, удобство использования. На основе анализа выявляются преимущества и недостатки предлагаемых решений, делаются выводы об актуальности разработки своего расширения, приводятся возможные идеи, рекомендации для разработки своего продукта, с учетом современных возможностей и технологий.

Ключевые слова: OCR, оптическое распознавание символов, браузерные расширения, Tesseract, Chromium, обработка изображений, распознавание текста, веб-технологии, автоматизация извлечения текста, цифровизация.

COMPARATIVE ANALYSIS OF OCR EXTENSIONS FOR CHROMIUM BROWSERS

Kim D.V.

Master's student

Kaluga State University named after K. E. Tsiolkovsky

Kaluga, Russia

Kuznetsova V.I.

PhD, Associate Professor

Kaluga State University named after K.E. Tsiolkovsky

Kaluga, Russia

Abstract

The article addresses the problem of extracting text from images and videos and aims to explore the OCR browser extension market. An analysis of existing offerings allows us to identify criteria for comparison: accuracy, speed, language support, and user-friendly interface. Based on the analysis, the advantages and disadvantages of the proposed solutions are identified, conclusions are drawn about the relevance of developing your own extension, and possible ideas and recommendations are provided for developing your own product, taking into account modern capabilities and technologies.

Keywords: OCR, optical character recognition, browser extensions, Tesseract, Chromium, image processing, text recognition, web technologies, automated text extraction, digitalization.

В связи с массовой цифровизацией в современном мире [10] почти вся текстовая информация переводится в цифровой формат, однако значительное количество данной информации часто хранится в виде изображений, PDF-Дневник науки | www.dnevnikaui.ru | СМН Эл № ФС 77-68405 ISSN 2541-8327

файлов и видеоконтента, из которых невозможно извлечь текст стандартными способами. Это обуславливает необходимость в использовании специализированных инструментов OCR, позволяющих преобразовать текст на изображениях в доступный для копирования формат.

Существует множество доступных для установки решений, но одними из наиболее удобных являются встроенные в браузеры семейства Chromium расширения - небольшие программы, которые выполняются внутри браузера и помогают расширить возможности сайтов [9]. С их помощью, такие браузеры как Google Chrome, Microsoft Edge, Opera, Яндекс Браузер становятся полезным и быстрым инструментом, позволяющим выделять области на странице, копировать с них текст, переводить его или редактировать.

Целью данной – провести сравнительный анализ OCR-расширений для браузеров на базе Chromium, выявить их функциональные и технические особенности, сформулировать рекомендации. Для достижения поставленной цели необходимо решить следующие задачи: рассмотреть особенности OCR как технологии распознавания текста; определить критерии сравнения OCR-расширений; выбрать актуальные решения для анализа; изучить их функции, преимущества и недостатки; сопоставить выбранные инструменты и сформулировать рекомендации для собственной разработки.

Для решения первой задачи сформулируем определение понятия «оптическое распознавание (OCR)» как процесс, который позволяет преобразовать изображения рукописного, машинописного или печатного текста в текстовые данные, используемые в компьютерах [3]. Данный процесс состоит из предварительной обработки текста, распознавания символов и обработки итогового результата. Существует два метода реализации – локально на устройстве пользователя или с помощью облачных сервисов. Преимуществом локальных решений является высокая конфиденциальность копируемой информации, так как текст не отправляется на внешние сервера. В свою

очередь, облачные сервисы обладают более высокими мощностями и функционалом.

Для проведения сравнительного анализа были выбраны наиболее важные в работе среднестатистического пользователя критерии, а именно: уровень удобства использования; способ обработки данных; наличие дополнительных функций; конфиденциальность данных.

Также необходимо определить: выполняется ли OCR локально на компьютере или через внешний сервис; какой алгоритм реализуется при работе с программой; нужно ли сохранять изображение или текст выделяется и копируется непосредственно на странице; может ли расширение работать с различными форматами, а не только со статичными картинками.

Наличие таких дополнительных функций как автоматическое определение ориентации текста, отображение текста поверх изначального изображения и перевод с иностранных языков помогает оценить сложность инструмента и его насыщенность. При этом важно, чтобы конечный пользователь был способен интуитивно понять и реализовать все предоставляемые возможности.

Для сравнительного анализа были выбраны пять OCR-расширений: Project Naptha; Copyfish Free OCR Software; OCR - Image Reader; Copy Text from Picture и Image Reader (OCR). Выбранные программные продукты позволяют провести содержательное сравнение, так как предоставляют разные подходы к реализации OCR в браузере.

Copyfish Free OCR Software – представляет собой стандартный вариант OCR-расширения. Пользователь выделяет область экрана вручную, после чего выполняется распознавание текста. Данный метод универсален и позволяет применять его ко всем медиа форматам [3], обеспечивает высокую скорость и точность. Также программа имеет дополнительные функции, такие как перевод текста и наложение его поверх изображений. Ядром инструмента является внешний OCR сервис [7], что исключает возможность автономной работы. Дневник науки | www.dnevnika.ru | СМИ Эл № ФС 77-68405 ISSN 2541-8327

OCR - Image Reader является локальным OCR-инструментом, распознающим текст на выделенной пользователем области. Благодаря использованию движка оптического распознавания Tesseract, повышается приватность при работе с конфиденциальными данными без необходимости обращения к внешним сервисам [6]. Дополнительные функции, такие как перевод текста и наложение поверх изображения отсутствуют, имеется возможность более тонкой настройки: выбор точности распознавания, автоматическое закрытие окна программы. К минусам можно отнести ручное изменение языка распознавания.

Project Naptha - представляет собой локальный инструмент, ориентированный на непосредственную работу с текстом внутри изображений на веб-страницах [8]. Отличительной особенностью является отсутствие окна захвата – картинка выделяется вся сразу, благодаря чему пользователь может взаимодействовать с изображением так, будто это обычный текст. Данный метод удобен в пределах страницы, но менее универсален по сравнению с расширениями, основанными на ручном выборе области. Кроме того, работа расширения в значительной степени зависит от характера изображения, размещённого на странице, что может ограничивать стабильность результата.

Copy Text from Picture – более простой и удобный вариант OCR-расширения, рассчитанный на повседневное пользование [1]. Имеет удобный и понятный интерфейс с минимумом настроек, что создает невысокий порог входа для любого пользователя, но не дает каких-либо дополнительных возможностей.

Image Reader (OCR) тяжелое расширение, открывающееся в отдельном окне [5]. Все изображения необходимо предварительно сохранить на устройстве, после чего загрузить их в программу. Отличается высокой скоростью работы, множеством настроек и возможностью скачать просканированный текст в формате .txt.

Сопоставление выбранных расширений показывает, что они различаются прежде всего по модели взаимодействия с пользователем. Project Naptha ориентирован на работу непосредственно с изображениями внутри веб-страницы, Image Reader (OCR) лучше всего работает с файлами, уже скачанными на компьютер, тогда как Copyfish, OCR - Image Reader и Copy Text from Picture используют более универсальный сценарий ручного выбора области. Это делает последние решения более удобными для обработки скриншотов, PDF-фрагментов и произвольных участков интерфейса.

По архитектуре обработки данных также есть отличия, наиболее популярным является локальная обработка изображений через OCR движки, такие как Tesseract, что обеспечивает высокую автономность и конфиденциальность, но ограничивает возможный функционал, который намного шире при использовании облачных решений, таких как Copyfish.

В результате перед пользователем встает выбор: использовать простые для понимания и удобные локальные решения или облачные решения с множеством дополнительных возможностей, но более сложным интерфейсом и ограниченной автономностью. Это позволяет сделать вывод о том, что в данный момент на рынке отсутствует универсальное решение, сочетающее оба подхода и обеспечивающее простоту освоения, конфиденциальность, полноту функционала и удобство работы.

Результаты проведенного сравнительного анализа OCR-расширений для браузеров семейства Chromium представлены в таблице 1.

Таблица 1 – Сравнительный анализ OCR-расширений для браузеров семейства Chromium

Расширение	Конфиденциальность данных	Удобство использования	Способ обработки данных	Дополнительные функции
Project Naptha	Высокая	Среднее + работа прямо на странице - мало гибкости	Локальный OCR	Отсутствуют

Расширение	Конфиденциальность данных	Удобство использования	Способ обработки данных	Дополнительные функции
Copyfish	Есть риски	Высокое + выделение области + автоматический выбор языка	Облачный OCR сервис	Перевод, наложение текста поверх изображения
OCR - Image Reader	Высокая	Среднее + выделение области - ручной выбор языка	Локальный OCR	Выбор точности распознавания, таймер закрытия окна программы
Copy Text from Picture	Высокая	Высокое + выделение области + минималистичный интерфейс - ручной выбор языка	Локальный OCR	Отсутствуют
Image Reader (OCR)	Высокая	Низкое - загрузка изображений из локального хранилища - ручной выбор языка	Локальный OCR	Скачивание просканированного текста в формате txt.

Таким образом, проведённое исследование показало, что OCR-расширения для браузеров на базе Chromium не образуют однородную группу решений. Их различия проявляются в архитектуре обработки данных, способе взаимодействия с пользователем, функциональной насыщенности и степени автономности.

Было обнаружено, что часть расширений основано на локальном распознавании текста, что обеспечивает более высокий уровень конфиденциальности и независимость от сетевого соединения. Другие решения опираются на облачные методы обработки, что расширяет их функциональные возможности и может положительно сказываться на качестве OCR в более сложных сценариях.

Установлено, что одни расширения более удобны при работе с изображениями, встроенными непосредственно в веб-страницы, другие при работе со скачанными на устройство файлами, тогда как третьи подходят для распознавания текста с произвольных областей экрана, PDF-файлов и других типов контента.

Можно сделать вывод, что на современном этапе не существует OCR-расширения, которое в полной мере сочетало бы простой пользовательский интерфейс, автономную локальную обработку, поддержку сложной структуры документов и универсальность практического применения. Это подтверждает целесообразность разработки собственного OCR-сканера изображений.

С нашей точки зрения, OCR-сканер следует разрабатывать по модульному принципу, используя локальный OCR в качестве базового режима и облачный OCR как дополнительный механизм для сложных изображений и документов. Также целесообразно реализовать простой пользовательский сценарий, включающий загрузку изображения или выбор области, запуск распознавания и получение редактируемого текстового результата. Дополнительно следует предусмотреть указание режима обработки данных, поддержку дальнейшей работы с распознанным текстом и возможность последующего расширения системы за счёт новых OCR-движков и функций обработки документов.

Библиографический список

1. Copy Text from Picture [Электронный ресурс] // Chrome Web Store. — URL: <https://chromewebstore.google.com/detail/copy-text-from-picture/lejhbacckkdpjnllmefhlinigbngpibl> (дата обращения: 22.03.2026).
2. Copyfish Free OCR Software [Электронный ресурс] // Chrome Web Store. — URL: <https://chromewebstore.google.com/detail/copyfish-%F0%9F%90%9F-free-ocr-soft/eenjdnjldapjajjofmldgmkjaienebbj> (дата обращения: 22.03.2026).
3. Gribomont, I. OCR with Google Vision API and Tesseract [Электронный ресурс] // Programming Historian. — 2023. — URL: <https://programminghistorian.org/en/lessons/ocr-with-google-vision-and-tesseract> (дата обращения: 22.03.2026).
4. How to use Copyfish [Электронный ресурс] // OCR.space Docs. — URL: <https://ocr.space/copyfish/docs> (дата обращения: 22.03.2026).

5. Image Reader (OCR) [Электронный ресурс] // Chrome Web Store. — URL: <https://chromewebstore.google.com/detail/image-reader-ocr/cakcfocedphbadddjpaiejbkhflfbhmf> (дата обращения: 22.03.2026).
6. OCR - Image Reader [Электронный ресурс] // Chrome Web Store. — URL: <https://chromewebstore.google.com/detail/ocr-image-reader/bhbhjkcoghibhibegcmbomkbackpdbo> (дата обращения: 22.03.2026).
7. OCR API [Электронный ресурс] // OCR.space. — URL: <https://ocr.space/ocrapi> (дата обращения: 22.03.2026).
8. Project Naptha [Электронный ресурс] // Chrome Web Store. — URL: <https://chromewebstore.google.com/detail/project-naptha/molncoemjfmrgdkbdlbjmhlcgniigdnf> (дата обращения: 22.03.2026).
9. What are extensions? [Электронный ресурс] // Chrome for Developers. — URL: <https://developer.chrome.com/docs/extensions/mv2/overview> (дата обращения: 22.03.2026).
10. Горелов, Н. А. Основы цифровой трансформации общества : учебник для вузов / Н. А. Горелов, О. Н. Кораблева. — 2-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2024. — 337 с. — ISBN 978-5-534-18432-7. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/535000> (дата обращения: 22.03.2026).