

АЛГОРИТМЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В РАСПОЗНАВАНИИ РУКОПИСНОГО ТЕКСТА В ДОКУМЕНТАЦИИ

Аннотация. В работе проведен анализ наиболее распространенных алгоритмов автоматического распознавания рукописного текста. Поскольку множество программных решений, широко представленных на рынке, не способны в полной мере распознать рукописный текст в офлайн-режиме, разработка системы, позволяющей это сделать, остается актуальной задачей.

Ключевые слова: искусственный интеллект, рукописные данные, документация, алгоритмы распознавания текста.

В настоящее время сканирование и сохранение в памяти компьютера текста с твердого носителя является решенной задачей. Это существенно облегчает задачу хранения рукописных и печатных текстов и предоставления доступа к ним различных пользователей. Однако, полученный в результате сканирования текст хранится в памяти компьютера в виде изображения, что делает работу с ним весьма сложной: затруднено ориентирование, практически невозможны редактирование, форматирование и поиск по тексту. Для решения этих задач необходимо провести процесс распознавания текста на изображении с созданием файла в том или ином текстовом формате [Святская].

Распознавание и оцифровка рукописного текста могут быть полезными во многих случаях. Например, это может быть полезно для автоматического перевода рукописных заметок в электронный вид, для архивирования документов или для удобного поиска и редактирования текста.

Цель работы заключается в повышении эффективности распознавания рукописного текста в документообороте посредством применения алгоритмов искусственного интеллекта.

Задачи:

1. Изучить основные принципы работы алгоритмов искусственного интеллекта, применяемых в распознавании рукописного текста.

¹ Научный руководитель: С. В. Ляхов, кандидат технических наук, доцент, заведующий кафедрой РГППУ.

2. Исследовать существующие методы распознавания рукописного текста и их применимость к документации.

3. Провести анализ преимуществ и недостатков различных подходов к распознаванию рукописного текста.

Распознавание может применяться в таких областях как [Огарок]:
внутренний документооборот на предприятии — распознавание определенных полей документов;

бумажная корреспонденция — распознавание ответов экзаменуемого в рамках какого-либо контрольного мероприятия;

распознавание конспектов, рукописных записей и заметок;

распознавание исторических документов.

Такие технологии могут значительно упростить работу с документами и делать их более доступными для использования в цифровом формате. К сожалению, программы, которые бы достаточно точно распознавали рукописный текст — отсутствуют. Для этого нужен искусственный интеллект. Поэтому, анкеты, заполненные вручную (например, клиентами в торговой точке, партнерами на выставке, кандидатами на работу в отделе кадров) — приходится вводить самостоятельно.

Работа распознавания рукописного текста носит название HWR (handwriting recognition). Существуют два различных класса задач HWR:

онлайн-распознавание — распознавание текста ведётся параллельно с вводом текста;

оффлайн-распознавание — распознавание текста ведётся на уже синтезированном изображении.

При онлайн-распознавании процесс формирования изображения текста и процесс его ввода в систему распознавания совмещены, что позволяет системе отслеживать процесс начертания символов. Это даёт возможность получать помимо графической информации ещё и информацию о структуре входных изображений, например, о направлении и скорости движения пера или о его нажиме при написании символа. На данный момент онлайн-системы распознавания широко используются на планшетных персональных компьютерах (ПК).

В оффлайн-распознавании, в отличие от предыдущей, система доступна только графическая информация. В этом случае, процесс распознавания базируется на меньшем количестве показателей и осложняет получение качественного результата. Неполный список проблем, типичных при распознавании рукописного текста оффлайн в общем случае, включает в себя [Козлов]:

высокая вариативность начертания символов — по размеру, наклону, набору составных частей, связям между ними и др.;

орфографические ошибки в тексте;
специфические особенности начертания, не позволяющие уверенно разделять символы;
пересечение элементов текста, наложение частей текста друг на друга;
помарки, кляксы, исправления, дефекты носителя (бумаги), а также артефакты, возникающие при сканировании;
не параллельность строк текста;
прочие проблемы.

Стандартное оффлайн-распознавание рукописного текста проводится по следующей схеме [Козлов]:

предобработка изображения, выделение области интереса;
сегментация и нормализация текста из области интереса;
распознавание сегментированного текста тем или иным методом.

Существует несколько популярных методов и подходов к распознаванию рукописного текста, таких как сверточные нейронные сети, рекуррентные нейронные сети, метод опорных векторов и другие.

Каждый из этих методов имеет свои преимущества и недостатки. Сверточные нейронные сети или CNN (от англ. Convolutional neural network) обычно используются для извлечения признаков из изображений, что делает их хорошим выбором для распознавания текста на изображениях с рукописным текстом.

Для распознавания рукописного текста с помощью CNN сначала изображения с символами преобразуются в черно-белые изображения. Затем нейронная сеть обрабатывает изображение, используя встроенные фильтры для выделения особенностей. После этого выдаются предсказания для каждого символа на изображении.

Обучение CNN для распознавания рукописного текста требует большого объема размеченных данных — т.е. изображений символов с правильными метками. Процесс обучения состоит в подаче изображений на вход нейронной сети, сравнении предсказаний с правильными метками и корректировке весов сети для улучшения точности предсказаний.

Рекуррентные нейронные сети или RNN (от англ. Recurrent neural network), в свою очередь, хорошо подходят для работы с последовательными данными, такими как последовательность символов в тексте. В распознавании рукописного текста, RNN могут быть использованы для последовательного анализа каждого пикселя изображения и предсказания символов текста. RNN могут быть обучены на большом наборе рукописных данных для обучения и дальнейшего распознавания текста на новых изображениях.

Для использования RNN подготавливаются обучающие данные в виде изображений рукописного текста и их соответствующих

меток (текстового представления). Далее изображения рукописного текста преобразуются в числовое представление, например, чтобы получить признаки изображения.

Создается RNN модель, которая будет обучаться на данных изображениях и их метках. RNN имеет способность запоминать предыдущие состояния и использовать их для предсказания следующего символа в тексте.

В свою очередь метод Опорных Векторов или SVM (от англ. Support Vector Machines) — это линейный алгоритм, используемый в задачах классификации и регрессии. Он работает, основываясь на концепции разделения данных на классы путем построения гиперплоскостей в признаковом пространстве.

Для распознавания рукописного текста с помощью SVM на первом этапе необходимо подготовить обучающий набор данных.

Затем данные изображения символов преобразуются в признаковое представление. Это может быть, например, представление в виде векторов пикселей изображений. Далее проводится обучение модели SVM, в результате чего строится гиперплоскость, разделяющая данные на различные классы символов.

После обучения модели SVM может быть использована для предсказания класса символа для новых, неизвестных изображений рукописного текста. Модель делает предсказание на основе расстояния до разделяющей гиперплоскости, выбирая класс символа, который находится на той же стороне гиперплоскости, что и данное изображение.

Основными принципами работы алгоритмов искусственного интеллекта, применяемых в распознавании рукописного текста, являются требование значительных вычислительных ресурсов, наличие качественного набора данных для обучения. Также стоит учитывать возможность комбинирования различных методов для достижения лучших результатов.

Системы распознавания текста могут быть применены в областях, где требуется обработка большого количества рукописных документов (заполненные бланки, анкеты), например, внутренний документооборот, бумажная корреспонденция, работа с историческими документами и т.д.

Сверточные нейронные сети хорошо подходят для обнаружения и выделения важных признаков изображений. Рекуррентные нейронные сети способны работать с последовательными данными. Метод опорных векторов может быть успешно применен для распознавания рукописного текста, обладая высокой точностью и эффективностью.

Список источников и литературы:

Святская Л. О. Методика распознавания рукописного текста в оффлайн-режиме // Наука, техника и образование. 2016. № 12 (30). С. 78–79. URL: <https://cyberleninka.ru/article/n/metodika-raspoznavaniya-rukopisnogo-teksta-v-offlayn-rezhime/viewer> (дата обращения: 08.02.2024).

Козлов В. Д. Методы оффлайн-распознавания рукописного текста (обзор). М.: МГУ, 2014. 14 с. URL: <http://textolog-rgali.ru/userfiles/articles/article2/GK.pdf> (дата обращения: 10.04.2024).

Галич А. С., Огарок А. Л. Комплексный алгоритм распознавания рукописного текста // Sci-article. 2023. № 117. URL: <https://sci-article.ru/stat.php?i=1684143179> (дата обращения: 08.04.2024).

УДК [001.891:81>276.6:62]:[81>322.4:004.8]

А. М. Хомутова¹

Российский государственный профессионально-педагогический университет

ПРИМЕНЕНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В НАУЧНЫХ ИССЛЕДОВАНИЯХ

Аннотация. Данная статья рассматривает роль искусственного интеллекта (ИИ) в преодолении языковых барьеров в научных исследованиях. Особое внимание уделяется анализу потенциала нейронных сетей для автоматизации перевода научной литературы, а также исследуются способы повышения эффективности машинного перевода с учетом лингвистических особенностей научных текстов.

Ключевые слова: искусственный интеллект, нейронные сети, машинный перевод, научная литература, лингвистический анализ.

Системы машинного перевода, основанные на нейронных сетях, способны обрабатывать большие объемы текста, обучаться на данных и адаптироваться к специфике научной терминологии. Исследования показывают, что точность машинного перевода научных текстов может варьироваться в зависимости от предметной области и сложности текста [Озерова]. В некоторых случаях машинный перевод может быть вполне приемлемым для понимания общей идеи текста, но может содержать неточности в деталях и терминологии.

¹ Научный руководитель: С. В. Ляхов, кандидат технических наук, доцент, заведующий кафедрой РГППУ.