

УДК 004.93

**О. И. Захарова,
к.т.н., доцент, зам. зав. НИЛ ИИ**

**В.Д. Бербасов,
магистрант**

**ФГБОУ ВО «Поволжский государственный университет
телекоммуникаций и информатики»,
Россия, Самара**

**ИССЛЕДОВАНИЕ СУЩЕСТВУЮЩИХ МЕТОДОВ
АВТОМАТИЧЕСКОЙ КОРРЕКЦИИ ОШИБОК**

Аннотация. Технологии распознавания текста и исправления ошибок играют ключевую роль в современном мире, охватывая широкий спектр областей, от оцифровки документов до преобразования текста в речь и услуг переводчика. С развитием инноваций и решением текущих проблем можно ожидать расширения возможностей и повышения эффективности систем распознавания текста и исправления ошибок в будущем.

Ключевые слова. Распознавание текста, исправление ошибок, оцифровка документов, преобразование текста в речь, многоязычные возможности, распознавание рукописного текста, конфиденциальность данных, обработка в реальном времени

O. I. Zakharova,

**Ph.D., Associate Professor, Deputy Head of the Laboratory of Artificial
Intelligence**

**V.D. Berbasov,
Master's student**

Samara, Russia

**RESEARCH ON EXISTING METHODS OF AUTOMATIC ERROR
CORRECTION**

Abstract: Text recognition and error correction technologies play a crucial role in today's world, spanning a wide range of areas from document digitization to text-to-speech conversion and translation services. With the development of innovations and the resolution of current issues, the expansion of capabilities and the enhancement of efficiency of text recognition and error correction systems can be expected in the future.

Keywords: Text recognition, error correction, document digitization, text-to-speech conversion, multilingual capabilities, handwriting recognition, data confidentiality, real-time processing

Введение

Оптическое распознавание, широко известное как OCR (оптическое распознавание символов), служит краеугольным камнем для преобразования печатного или рукописного текста из изображений или документов в цифровые форматы. Его приложения охватывают все - от сканирования документов и автоматического ввода данных до извлечения текста из изображений. С растущим внедрением технологии распознавания текста становится очевидной необходимость в надежных механизмах исправления ошибок.

Распознавание текста коренным образом изменило наше взаимодействие с печатными и рукописными материалами, упростив задачи ввода данных и позволив оцифровать обширные текстовые архивы. Однако даже самые продвинутые системы распознавания текста не застрахованы от ошибок.

Только в последние десятилетия были достигнуты значительные прорывы благодаря развитию методов машинного обучения и компьютерного зрения. Современные системы распознавания текста используют сочетание распознавания образов, нейронных сетей и алгоритмов глубокого обучения, таких как сверточные нейронные сети (CNN), для распознавания и интерпретации текста. Эти достижения

значительно повысили точность и устойчивость технологии распознавания текста.

Системы распознавания текста используют множество методов и алгоритмов для извлечения текста из изображений или документов, которые обычно делятся на две категории: основанные на функциях и основанные на глубоком обучении.

Методы, основанные на признаках, предполагают извлечение определенных характеристик или отличительных черт из входного изображения, таких как края и углы, что помогает идентифицировать символы и слова. Классические алгоритмы машинного обучения, такие как машины опорных векторов (SVM) и скрытые марковские модели (HMM), часто дополняют распознавание на основе признаков.

Напротив, в последние годы все большее распространение получили методы, основанные на глубоком обучении. Сверточные нейронные сети (CNN) продемонстрировали замечательное мастерство в распознавании текста внутри изображений. Эти сети обладают способностью извлекать иерархические объекты непосредственно из пиксельных данных, что делает их исключительно хорошо подходящими для задач распознавания текста.

Классификация ошибок при распознавании

Ошибки при оптическом распознавании можно разделить на различные типы, каждый из которых представляет свой собственный набор проблем и последствий для исправления ошибок:

Ошибки распознавания символов: Эти ошибки проявляются, когда отдельные символы неправильно интерпретируются или идентифицируются некорректно. Такие факторы, как вариации шрифта, размер символа и перекося текста, способствуют ошибкам распознавания символов.

Ошибки шума и искажения: Шум изображения, низкое качество изображения и искажения, возникающие в результате сканирования или захвата изображений устройствами, вносят ошибки в процесс распознавания. Для устранения этих ошибок требуются методы удаления шума и улучшения изображения.

Ошибки компоновки и форматирования: В контексте документов с несколькими столбцами или сложных макетов возникают ошибки при определении правильной последовательности и размещения распознанного текста. Сохранение точной структуры документа имеет решающее значение для исправления ошибок в таких сценариях.

Языковые и контекстные ошибки: Понимание контекста и семантики текста представляет собой серьезную проблему при распознавании текста. Ошибки в распознавании языка, устранении неоднозначности смысла слов и контекстуальном понимании приводят к ошибочным интерпретациям.

Существующие методы исправления ошибок

Жизненно важно обеспечить надежность и точность систем распознавания текста. Для исправления ошибок при оптическом распознавании было разработано несколько методов и техник, каждый из которых обладает своими уникальными преимуществами и ограничениями.

Методы постобработки: Распространенный подход включает методы постобработки, которые исправляют ошибки после извлечения текста. В этих методах часто используются алгоритмы проверки орфографии для обнаружения и исправления ошибок распознавания символов. Кроме того, инструменты коррекции грамматики повышают качество распознаваемого текста за счет тщательного изучения контекста и синтаксиса.

Подходы, основанные на машинном обучении: Использование моделей машинного обучения для исправления ошибок в последние годы получило значительное распространение. Эти модели проходят обучение с

использованием обширных наборов данных, содержащих текст и ошибки, что позволяет им выявлять закономерности и несоответствия в распознанном тексте. Многообещающие результаты были получены благодаря внедрению рекуррентных нейронных сетей (RNNS) и моделей "последовательность к последовательности" в задачах исправления ошибок.

Циклы обратной связи: Внедрение механизмов обратной связи в системах распознавания текста представляет собой еще один способ исправления ошибок. Эти механизмы извлекают информацию из предыдущих ошибок распознавания, используя эту обратную связь для адаптации и повышения точности распознавания с течением времени. Постоянно совершенствуя модели распознавания, контуры обратной связи повышают производительность системы.

Недавние успехи в исправлении ошибок при оптическом распознавании были поистине замечательными, чему способствовала доступность обширных наборов данных и возросшие вычислительные возможности современного аппаратного обеспечения. Эти усовершенствования привели к существенному повышению точности распознавания и надежности.

Заслуживающая внимания тенденция связана с внедрением методов обработки естественного языка (NLP) в системы распознавания текста. Модели НЛП, такие как архитектуры на основе трансформаторов, произвели революцию в контекстуальном понимании текста. Эти модели расшифровывают коннотации и намерения, стоящие за словами и фразами, эффективно сокращая количество языковых и контекстных ошибок при распознавании текста.

Состязательное обучение и методы увеличения объема данных также сыграли ключевую роль в повышении устойчивости систем распознавания текста. Состязательное обучение включает в себя подготовку моделей

распознавания с использованием составительных примеров, специально созданных для того, чтобы проверить устойчивость системы к шуму и искажениям. Методы увеличения объема данных искусственно вносят изменения в обучающие данные, имитируя условия реального мира.

Возможные области применения рассмотренных технологий

Области применения распознавания текста и исправления ошибок обширны и постоянно расширяются в tandem с технологическим развитием. Известные области применения включают:

Оцифровка документов: Технология распознавания текста является основой для преобразования обширных хранилищ бумажных документов в цифровые форматы, что приносит пользу библиотекам, архивам и организациям, работающим с обширными коллекциями документов.

Извлечение данных: В различных областях, таких как финансы, здравоохранение и администрирование, распознавание текста играет ключевую роль в автоматизации работы по вводу данных. Счета-фактуры, бланки, квитанции и другие структурированные документы могут быть эффективно обработаны.

Преобразование текста в речь: Распознавание текста играет ключевую роль в облегчении доступа к печатным или рукописным материалам для людей с нарушениями зрения. Преобразование текста в синтезированную речь значительно повышает доступность информации.

Услуги переводчика: Многоязычные возможности OCR расширяют возможности перевода печатного текста на различные языки, облегчая глобальную коммуникацию и понимание.

Несмотря на значительные успехи в оптическом распознавании и исправлении ошибок, сохраняется целый ряд проблем, подчеркивающих необходимость их решения, поскольку спрос на точное распознавание продолжает расти.

Распознавание рукописного текста: Точное распознавание рукописного текста остается сложной задачей из-за различий в стилях письма и отсутствия стандартизированных шрифтов. Исследовательские усилия должны быть сосредоточены на улучшении распознавания рукописного текста.

Распознавание нелатинских шрифтов: В то время как системы распознавания текста в основном разрабатывались для латинских шрифтов, точность распознавания нелатинских шрифтов может быть значительно ниже. Расширение возможностей распознавания текста для охвата более широкого спектра сценариев имеет решающее значение для глобального применения.

Обработка в реальном времени: Распознавание текста в реальном времени, жизненно важное для таких приложений, как мобильные приложения для сканирования и дополненная реальность, требует повышенного внимания. Будущие инновации должны быть направлены на расширение возможностей обработки данных в режиме реального времени, чтобы соответствовать ожиданиям пользователей.

Проблемы конфиденциальности: С распространением технологии распознавания текста возникли проблемы, связанные с конфиденциальностью и безопасностью данных. Этические соображения, связанные с обработкой конфиденциальной информации, полученной с помощью распознавания текста, требуют тщательного обдумывания.

Заключение

Следует отметить, что автоматическая коррекция ошибок при оптическом распознавании незаменима в мире, где оцифровка и доступность данных имеют первостепенное значение. По мере развития технологий будущее распознавания текста и исправления ошибок обещает дальнейшую оптимизацию управления данными и повышение доступности.

Область оптического распознавания динамична и постоянно развивается. Исследователи и практики готовы преодолевать существующие препятствия и использовать появляющиеся возможности, продвигая технологию распознавания текста к большей точности, универсальности и доступности.

Список источников

1. Доронина А.В., Захарова О.И. Нейросетевые алгоритмы и их виды. Анализ существующих нейросетевых алгоритмов. Роль математики. Научный аспект. 2023. Т. 18. № 6. С. 2261-2267.

2. Скрытые Марковские модели [Электронный ресурс] Режим доступа: https://neerc.ifmo.ru/wiki/index.php?title=%D0%A1%D0%BA%D1%80%D1%8B%D1%82%D1%8B%D0%B5_%D0%9C%D0%B0%D1%80%D0%BA%D0%BE%D0%B2%D1%81%D0%BA%D0%B8%D0%B5_%D0%BC%D0%BE%D0%B4%D0%B5%D0%BB%D0%B8, свободный. – Загл. с экрана.

3. Метод опорных векторов (SVM) [Электронный ресурс] Режим доступа: [https://neerc.ifmo.ru/wiki/index.php?title=%D0%9C%D0%B5%D1%82%D0%BE%D0%B4_%D0%BE%D0%BF%D0%BE%D1%80%D0%BD%D1%8B%D1%85_%D0%B2%D0%B5%D0%BA%D1%82%D0%BE%D1%80%D0%BE%D0%B2_\(SVM\)](https://neerc.ifmo.ru/wiki/index.php?title=%D0%9C%D0%B5%D1%82%D0%BE%D0%B4_%D0%BE%D0%BF%D0%BE%D1%80%D0%BD%D1%8B%D1%85_%D0%B2%D0%B5%D0%BA%D1%82%D0%BE%D1%80%D0%BE%D0%B2_(SVM)), свободный. – Загл. с экрана.

4. Что такое обработка естественного языка (NLP)? [Электронный ресурс] Режим доступа: <https://aws.amazon.com/ru/what-is/nlp/>, свободный. – Загл. с экрана.