

(не только стандартных форматов офисных программ), высокая способность к интеграции с иными корпоративными системами.

Таким образом, данные технологии стали реализовываться и в системах электронного документооборота, что позволило СЭД эффективнее и быстрее подстраиваться под нужды быстрорастущих организаций. На сегодняшний день большинство программных продуктов, представленных на отечественном и зарубежном рынке, включают в себя технологии потоков работ, управления цифровыми активами. Многие современные системы электронного документооборота позволяют управлять бизнес-процессами организации, а облачные технологии и технологии искусственного интеллекта начинают использоваться в данном виде ПО все чаще. Например, такие системы, как ELMA BPM, ТЕЗИС включают в себя модуль управления бизнес-процессами. К классу ECM-систем можно отнести системы EdocLib, Электронное Дело, SMBusiness и т. д. Примером CSP-платформ является платформа DocsHouse (разработка ЛАНИТ), переходят от концепции ECM к концепции CSP разработчики систем Docsvision и Directum (а также ее редакция RX) Bel. В СЭД DIRECTUM Bel присутствует также возможность установки модуля Ario One – системы искусственного интеллекта для обработки документов, позволяющая автоматически распределять корреспонденцию по видам и журналам регистрации, заносить реквизиты в карточку документа и т. д.

Список использованных источников

1. Kitta, T. SharePoint Server 2010 Enterprise Content Management / T. Kitta, B. Grego, R. Houberg. – Washington : John Wiley&Sons, 2011. – 480 p.
2. Сафронов, А. В. Компьютеризация управления плановой экономикой в СССР: проекты ученых и нужды практиков [Электронный ресурс]. / А. В. Сафронов // КиберЛенинка – научная электронная библиотека. – Режим доступа: <https://cyberleninka.ru/article/n/kompyuterizatsiya-upravleniya-planovoy-ekonomikoy-v-sssr-proekty-uchenyh-i-nuzhdy-praktikov>. – Дата доступа: 18.02.2024.
3. Woodbridge, M. The Death of ECM and Birth of Content Services / M. Woodbridge // Gartner Inc – a technological research and consulting firm [Electronic resource]. – Mode of access: <https://blogs.gartner.com/michael-woodbridge/the-death-of-ecm-and-birth-of-content-services/>. – Date of access: 18.02.2024.

УДК 004.8

МАШИННОЕ РАСПОЗНАВАНИЕ И ОБРАБОТКА ТЕКСТОВЫХ СИМВОЛОВ

Антонова Т.А., студ., Дунина Е.Б., к.т.н., доц.
*Витебский государственный технологический университет,
г. Витебск, Республика Беларусь*

Реферат. В статье рассмотрены способы и средства для машинного распознавания, а также обработки текстовых символов, приведён пример разработки программного средства для распознавания государственных номеров автомобилей.

Ключевые слова: искусственный интеллект, машинное обучение, компьютерное зрение, оптическое распознавание символов, датасет, аннотация, государственные номера автомобилей, Python, Tesseract, OpenCV, Ultralytics, YOLOv8, OCR, open-source.

Компьютерное зрение – это область искусственного интеллекта, связанная с анализом изображений и видео. Она включает в себя набор методов, которые наделяют компьютер способностью «видеть» и извлекать информацию из увиденного.

Системы состоят из фото- или видеокамеры и специализированного программного обеспечения, которое идентифицирует и классифицирует объекты. Они способны анализировать образы (фотографии, картинки, видео, штрих-коды), а также лица и эмоции.

Чтобы научить компьютер «видеть», используются технологии машинного обучения.

Оптическое распознавание текста – одна из частей компьютерного зрения, которая отвечает за считывание текста с изображений для дальнейшей машинной обработки. Данная технология

включает в себя обнаружение текстового содержимого на изображениях и перевод изображений текста в закодированный текст, который компьютер может легко понять. Изображение, содержащее текст, сканируется и анализируется, чтобы идентифицировать символы в нем. После идентификации символ преобразуется в текст.

Благодаря этой технологии человеку стали доступны многие возможности автоматизации. В данной статье будет рассмотрен вопрос распознавания номерных знаков автомобилей на КПП – контрольно-пропускных пунктах.

Распознавание номерного знака и государственного номера автомобилей на КПП требуется для автоматизации контроля пропускного режима при въезде транспортных средств на территорию предприятия.

Для осуществления данной цели используется нейронная модель YOLOv8 – это одна из наиболее известных моделей для обнаружения объектов на изображениях в реальном времени. Она основана на сверточных нейронных сетях и позволяет достичь высокой скорости обработки без ущерба точности. YOLO разделяет изображение на сетку ячеек, и каждая ячейка предсказывает границы и классы объектов, содержащихся внутри нее.

Для полноценной работы требуется произвести дообучение модели на датасете со многочисленными снимками машин с государственными номерами Республики Беларусь для того, чтобы модель имела возможность распознать номерной знак на изображении. Датасет был сформирован из порядка 10 000 изображений, включая различные ракурсы, углы наклона номерного знака. Подобные примеры нужны для того, чтобы модель могла работать вне «идеальных» условий и распознавать объекты корректно в большинстве случаев. Датасет для обучения моделей семейства YOLO обычно имеет характерную структуру (рис. 1): древовидная система хранения файлов (изображений и аннотаций к ним).

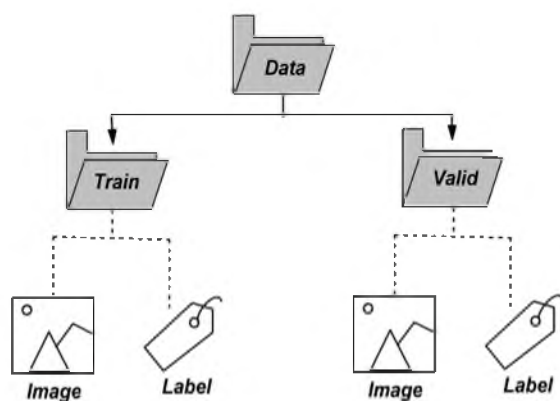


Рисунок 1 – Файловая структура датасета

Аннотации – это текстовые файлы, в которых записана информация о классе объекта распознавания и координаты центра этого объекта. Пример пары изображение-аннотация из датасета представлен на рисунке 2.

При обучении моделей нейронных сетей часто используется термин «эпоха» – один полноценный проход модели по заготовленному датасету. Чем больше эпох пройдет модель при обучении, тем больше раз датасет полностью будет использован. Важно подобрать количество эпох, чтобы достичь максимального результата за меньшее количество эпох.

Важно отметить, что для обучения чаще всего может потребоваться графический процессор, так оно пройдет быстрее. В зависимости от

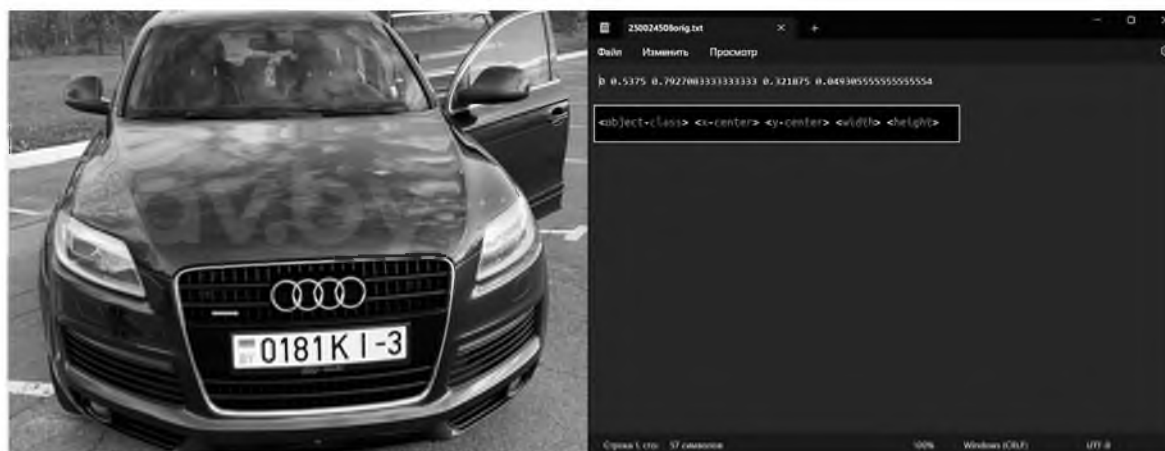


Рисунок 2 – Пример пары изображение-аннотация из датасета

характеристик видеокарты нужно подбирать настроечные параметры для обучения.

Таким образом, модель была успешно обучена. На выходе имеется ограниченная область с номерным знаком на изображении.

После того, как знак был выделен, производится OCR (Optical Character Recognition, оптическое распознавание текста) с помощью Tesseract и конвертация его в строковое значение.

Tesseract – это движок оптического распознавания символов от Google (OCR) с открытым исходным кодом, является самой популярной и качественной OCR-библиотекой.

Python-tesseract – это инструмент оптического распознавания символов (OCR) для python. Pytesseract может считывать все типы изображений.

OCR использует нейронные сети для поиска и распознавания текста на изображениях.

Tesseract ищет шаблоны в пикселях, буквах, словах и предложениях, использует двухэтапный подход, называемый адаптивным распознаванием. Требуется один проход по данным для распознавания символов, затем второй проход, чтобы заполнить любые символы, в распознавании которых он не был уверен, символами, которые, скорее всего, соответствуют данному слову или контексту предложения.

В данном случае логика состоит в том, чтобы вначале максимально подготовить изображение к OCR, применяя фильтры и масштабируя, выделить контуры номерного знака и текстовых символов, а затем применить метод конвертации изображения в текст.

За подготовку изображения отвечает библиотека OpenCV – это open-source библиотека компьютерного зрения, которая предназначена для анализа, классификации и обработки изображений.

Таким образом, результатом работы программных средств является распознанный номерной знак и символьная его интерпретация (рис. 3).

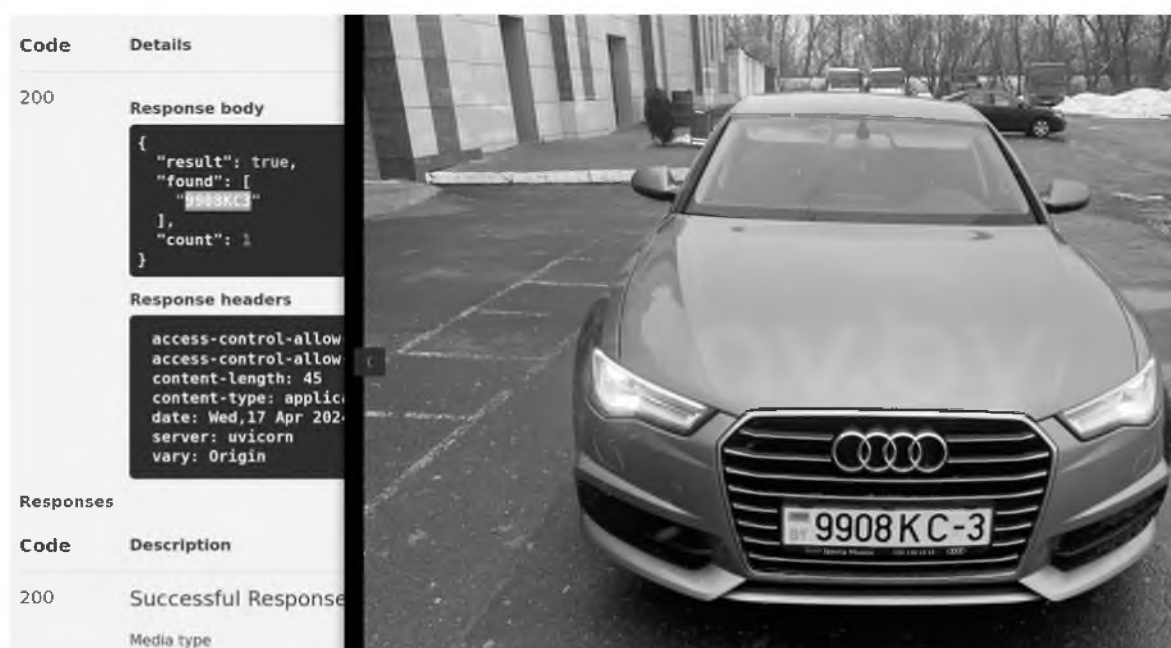


Рисунок 3 – Результат работы программных средств

Список используемых источников

1. Лабораторно-техническое направление развитие кафедры АТПП / А. А. Кузнецов [и др.]: материалы докладов 47 международной научно-технической конференции преподавателей и студентов / УО «ВГТУ». – Витебск, 2014. – С. 462–465.
2. Обучите YOLOv8 на пользовательском наборе данных [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/articles/714232/>. – Дата доступа: 02.05.2024.
3. Распознавание текста с помощью pytesseract [Электронный ресурс]. – Режим доступа: https://pikabu.ru/story/raspoznavanie_teksta_s_pomoshchyu_pytesseract_9713785. – Дата доступа: 02.05.2024.