



Столыпинский
вестник

Научная статья

Original article

УДК 004.414.2

**ВЫБОР СРЕДСТВ РЕАЛИЗАЦИИ ПРИ РАЗРАБОТКЕ
АВТОМАТИЗИРОВАННОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ С
ПРИМЕНЕНИЕМ ТЕХНОЛОГИЙ МАШИННОГО ЗРЕНИЯ**
CHOICE OF IMPLEMENTATION TOOLS IN DEVELOPING AUTOMATED
INFORMATION SYSTEM USING MACHINE VISION TECHNOLOGIES

Ветохин Валерий Викторович, канд. техн. наук, доцент кафедры компьютерных интеллектуальных технологий проектирования, Воронежский государственный технический университет, г.Воронеж

Ракитин Ян Евгеньевич, бакалавр, Воронежский государственный технический университет, г.Воронеж

Нестеренко Илья Александрович, бакалавр, Воронежский государственный технический университет, г.Воронеж

Vetokhin Valery Viktorovich, Cand. Sc. (Technical), Associate Professor of the Department of Computer Intelligent Design Technologies, Voronezh State Technical University, Voronezh, e-mail: daiolix@yandex.ru

Rakitin Yan Evgenyevich, Bachelor, Voronezh State Technical University, Voronezh, e-mail: y.rakitin0802@gmail.com

Nesterenko Ilya Alexandrovich, Bachelor, Voronezh State Technical University, Voronezh, e-mail:kitp@vorstu.ru

Аннотация

В данной статье идет речь о выборе программных компонентов при разработке автоматизированной информационной системы с применением технологий машинного зрения. Было отмечено, что выбор инструментов реализации на этапе проектирования программного продукта крайне важен и является определяющим в разработке. Для того, чтобы оптимально использовать имеющиеся технологии компьютерного зрения и оптического распознавания текста произведен анализ методов и построены выводы на основе полученных данных. Авторы отмечают, что при обработке изображения в режиме реального времени, когда границы объектов не явно различимы и стремительно меняются использование алгоритмов классического компьютерного зрения не целесообразно, намного эффективнее в задачах подобного рода показывают себя алгоритмы на основе машинного обучения.

Annotation

This article deals with the choice of software components in the development of an automated information system using machine vision technologies. It was noted that the choice of implementation tools at the design stage of a software product is extremely important and is decisive in development. In order to optimally use the available technologies of computer vision and optical text recognition, an analysis of the methods was made and conclusions were drawn based on the data obtained. The authors note that when processing an image in real time, when the boundaries of objects are not clearly distinguishable and rapidly changing, the use of classical computer vision algorithms is not advisable, algorithms based on machine learning are much more effective in tasks of this kind.

Ключевые слова: машинное обучение, машинное зрение, компьютерное зрение, оптического распознавание текста, трансферное обучение.

Key words: machine learning, machine vision, computer vision, optical text recognition, transfer learning.

Приступая к разработке автоматизированной информационной системы или же алгоритма с применением технологий машинного зрения необходимо понимать какие инструменты и решения следует использовать для наиболее эффективного функционирования конечного программного продукта. На примере разработки автоматизированной информационной системы по распознаванию регистрационных знаков автомобильных номеров рассмотрим некоторую часть процесса создания программы.

Задачу определения объектов на изображении решают методы компьютерного зрения, оно же машинное зрение. Изучив теоретический материал по данной теме необходимо построить определенные выводы о том, какие конкретно подходы стоит использовать. Алгоритмы компьютерного зрения можно разделить на две категории:

- «Классические» алгоритмы компьютерного зрения
- Алгоритмы с применением машинного обучения

Два подхода описанные выше разительно отличаются друг от друга. Методы классического компьютерного зрения показывают себя намного хуже, а порой и вовсе не реализуемы, в сравнении с методами с применением алгоритмов машинного обучения в тех случаях, когда перед программой стоит задача классификации множества объектов на изображении в режиме реального времени. Классические алгоритмы можно использовать, когда необходимо работать со статическими или короткими видеофрагментами, на которых границы объектов чётко отражены и выводы, которые должна предоставить программа будут заключаться на основе качественных характеристик изображения, таких как цвет, область цвета, идентичность точек на двух сравниваемых изображениях, а не в задачах, которые под силу интуитивно и быстро способен решать только человек. Основным преимуществом использование нейросетевых технологий является хорошая обобщающая способность, возможность использовать контекстный анализ и распознавать символ, основываясь на окружающие его символы. [1, с.1099] На такое способны и неуклонно растут в данном направлении алгоритмы на основе машинного обучения, ведь в процессе обучения нейросеть,

что структурно похожа на человеческий мозг, эмитирует предсказания и строит выводы на тех размеченных данных, на которые уже четко были выявлены объекты интереса. Следует обратить внимание на такой подход обучения, как трансферное обучение. Трансферное обучение — это подраздел машинного обучения, целью которого является применение знаний, полученные из одной задачи, к другой целевой задаче. [2] Раньше для того, чтобы заняться глубинным обучением, вы должны были иметь доступ к большому очищенному набору данных и самостоятельно разработать, и обучить эффективную модель. [3] С помощью описанного выше подхода к решению задач с применением нейросетевых технологий у разработчика пропадает необходимость собственноручно проектировать нейросетевую модель, что вовсе не является простой задачей. Такой подход позволяет сэкономить время, человеческие ресурсы, а порой и вовсе использовать оптимальную архитектуру. Для успешного трансферного обучения нужна модель, натренированная на наборе данных, схожем с целевым. Так, модель, обученная на изображениях лиц, не лучшим образом обобщается для задачи распознавания текста. Поэтому популярны модели обучены на больших и разнообразных наборах данных.[4] Построенная нейросетью нетривиальная зависимость позволяет заложенной в нее архитектурной конструкции отмечать такие объекты уже на других, «чистых» и неразмеченных изображениях. Но у алгоритмов на основе нейросетей есть и свои характеристики, что делают их различными: время обучения, количество слоев, количество весовых коэффициентов, математическое обоснование и другие. Для того, чтобы упростить работу по выбору макета нейросети можно воспользоваться интернет ресурсом «paperswithcode.com» на котором расположена информация по различным ключевым архитектурам популярных нейросетей, что значительно упрощает процесс выбора оптимальной технологии.

Задача распознавания текста на изображении, так же, как и методы решения задач компьютерного зрения можно разделить на две категории: распознавание с помощью метрик и распознавание с помощью нейросетей. В связи с возникновением новой волны популярности нейросетевых классификаторов они

стали чаще использоваться в исследовательских работах по распознаванию текста. [5, с.249] Зачастую, проектирование «с нуля» модели машинного обучения при решении задач подобного рода не уместно для того, чтобы корректно и с определенной частотой положительных, правильных выводов обнаружения и определения текста на изображении необходимо найти хорошую обучающую выборку, влияние на точность детекции будет крайне существенно зависеть от качества данных для обучения (их количества и качества) настроить весовые коэффициенты сети под конкретные типы изображений, произвести дополнительные работы по тестированию. И в дальнейшем аппаратные ресурсы системы будут задействованы во много большем объеме, чем если бы в работах по оптическому распознаванию текста были задействованы методы распознавания при помощи готовых открытых решений. Они не уступают по эффективности, но при этом их легче интегрировать в проект, такие решения задействуют меньше системных ресурсов и не требуют специализированного программного обеспечения, кроме самого языка программирования, на котором будет реализован конечный продукт.

Аппаратно-техническое обеспечение устанавливается заказчиком программного продукта исходя из системных характеристик оборудования, на котором в дальнейшем информационная система будет функционировать. С аппаратной точки зрения системы с применением технологий компьютерного зрения включают в себя средство захвата изображения, электронно-вычислительную машину и специализируемое программное обеспечение. [6, с.319] При разработке будут реализованы отдельные модули оптимизации работы программы в которых возможно без особых усилий в дальнейшем изменить объемы используемых системных характеристик, к примеру, в ходе выполнения работы было установлено, что использование всего объема видеопамати графической карты может привести к сбоям в работе системы и ее отказу, во избежание проблем подобного рода был разработан скрипт, ограничивающий использование видеопамати, в котором значение равное разрешенному для

использования объема видеопамати может быть изменено от одних системных характеристик к другим.

Приступим к проектированию автоматизированной информационной системы с применением технологий машинного зрения. Целью проектирования программного продукта, перед его практической реализацией, является определение внутренних свойств системы и детализация ее внешних свойств на основе выдвинутых в техническом задании требований к программному обеспечению. Ключевые особенности проектируемой информационной системы накладывают ограничения на бизнес-логику и структуры данных, используемые в конечном продукте.

В качестве примера будет использована предобученная нейросетевая модель от TensorFlow под названием SSD MobileNetV2 и система оптического распознавание текста EasyOCR. Для начала нужно определить принцип работы алгоритма детекции на основе SSD MobileNetV2 320*320, на вход которого информация поступает в виде изображения в одном формате и возвращает такой алгоритм на выходе тоже изображение с различными метками на нем. Алгоритм работы нейросетевой модели SSD, представлен на рисунке 1.

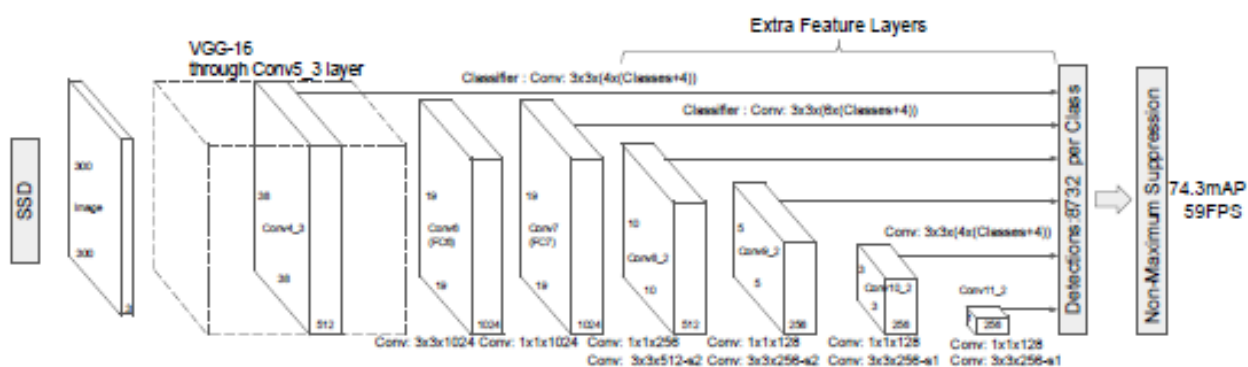


Рисунок 1 – принцип работы нейросети SSD MobileNetV2

Алгоритм оптического распознавания текста EasyOCR (рис 2) на вход принимает изображение, а возвращает текстовые данные.

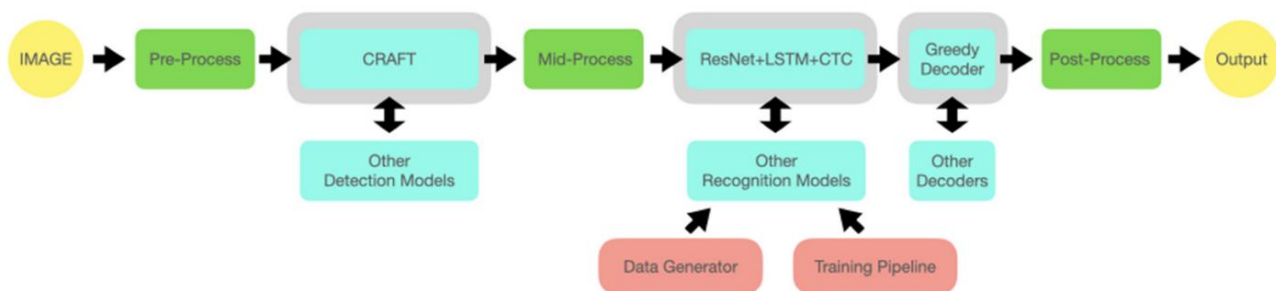


Рисунок 2 – принцип работы EasyOCR

Следует отметить, что при выполнении алгоритма определения региона с расположением автомобильных номеров на изображении метод вернет текстовые данные, обозначающие разметку объекта детекции. Эти данные несут в себе значения координат области с государственным регистрационным знаком автомобиля, на котором и будут производиться работы алгоритма оптического распознавания текста, то есть в дальнейшем, после выполнения алгоритма детекции необходимо в целях понижения затрачиваемых аппаратно-технических мощностей уменьшить исследуемое изображение, используя в последующих участках программы только ту область, в которой необходимо произвести, к примеру, оптическое распознавание текста.

В дальнейшем, получив на выходе регионы с различными текстовыми данными, полученными в следствии выполнения алгоритма с использованием методов EasyOCR, стало ясно, что на номерных знаках автомобиля кроме ключевой информации содержится информационный мусор, который необходимо отфильтровать и удалить, процесс создания алгоритма фильтрации описан в практической реализации программного продукта. Исходя из всего вышеперечисленного некоторые модули программы должны быть следующими:

- Алгоритм детекции на основе SSD MobileNetV2 320*320
- Алгоритм сохранения исследуемой области на изображении
- Алгоритм оптического распознавания текста на основе EasyOCR
- Алгоритм фильтрации текста в области ограничивающего прямоугольника

Литература

1. Илларионов А. А., Чернов А. А. ОБЗОР МЕТОДОВ РАСПОЗНАВАНИЯ ТЕКСТА //Аллея науки. – 2019. – Т. 2. – №. 5. – С. 1094-1098.
2. [Электронный ресурс] // Transfer Learning / Python-School. URL: [https://python-school.ru/wiki/transfer-learning/#:~:text=Transfer%20Learning%20\(трансферное%20обучение\)%20—,поэтому%20нет%20необходимости%20изобретать%20велосипед](https://python-school.ru/wiki/transfer-learning/#:~:text=Transfer%20Learning%20(трансферное%20обучение)%20—,поэтому%20нет%20необходимости%20изобретать%20велосипед). (дата обращения: 27.07.2022г)
3. [Электронный ресурс] // Трансферное обучение: почему deep learning стал доступнее / Академия Яндекса. URL: <https://academy.yandex.ru/posts/transfernnoe-obuchenie-pochemu-deep-learning-stal-dostupnee> (дата обращения: 27.07.2022г)
4. [Электронный ресурс] // Сорока.Е. Предобученные модели / robot_dreams. URL: <https://robotdreams.cc/blog/115-predobuchennye-modeli> (дата обращения: 27.07.2022г)
5. Спицын В. Г. и др. Применение вейвлет-преобразования Хаара, метода главных компонент и нейронных сетей для оптического распознавания символов на изображениях в присутствии импульсного шума //Компьютерная оптика. – 2016. – Т. 40. – №. 2. – С. 249-257.
6. Горячкин Б. С., Китов М. А. Компьютерное зрение //E-Scio. – 2020. – №. 9 (48). – С. 317-345.

Literature

1. Illarionov A. A., Chernov A. A. REVIEW OF TEXT RECOGNITION METHODS // Alley of Science. - 2019. - Vol. 2. - No. 5. - P. 1094-1098.
2. [Electronic resource] // Transfer Learning / Python-School. URL: <https://python-school.ru/wiki/transfer-learning/#:~:text = Transfer%20Learning%20> (date of application: 27.07.2022)
3. [Electronic resource] // Transfer learning: why deep learning has become more accessible / Yandex Academy. URL: <https://academy.yandex.ru/posts/transfernnoe-obuchenie-pochemu-deep-learning-stal-dostupnee> (Date of application: 07/27/2022)

4. [Electronic resource] // Soroka.E. Pre-trained models / robot_dreams. URL: <https://robotdreams.cc/blog/115-predobuchennye-modeli> (Date of application: 07/27/2022)
5. Spitsyn V. G. et al. Application of the Haar wavelet transform, the method of principal components and neural networks for optical character recognition in images in the presence of impulse noise // Computer Optics. - 2016. - Т. 40. - No. 2. - P. 249-257.
6. Goryachkin B.S., Kitov M.A. Computer vision //E-Scio. – 2020. – no. 9 (48). – P. 317-345.

© Ветохин В. В., Ракитин Я. Е., Нестеренко И. А., 2022 // Научный сетевой журнал «Столыпинский вестник», №4/2022.

Для цитирования: Ветохин В. В., Ракитин Я. Е., Нестеренко И. А. ВЫБОР СРЕДСТВ РЕАЛИЗАЦИИ ПРИ РАЗРАБОТКЕ АВТОМАТИЗИРОВАННОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ С ПРИМЕНЕНИЕМ ТЕХНОЛОГИЙ МАШИННОГО ЗРЕНИЯ // Научный сетевой журнал «Столыпинский вестник», №4/2022.