

Бадика Егор Максимович, Студент,
Санкт-Петербургский Государственный Университет
Аэрокосмического Приборостроения, Санкт-Петербург

Кузьменко Владимир Павлович, Кандидат технических наук,
Санкт-Петербургский Государственный Университет
Аэрокосмического Приборостроения, Санкт-Петербург

МОДЕЛЬ ИНИЦИАЛИЗАЦИИ ПРОМЫШЛЕННЫХ РОБОТОВ С ПОМОЩЬЮ ОБНАРУЖЕНИЯ ОБЪЕКТОВ НА ОСНОВЕ ГЛУБОКОГО ОБУЧЕНИЯ

Аннотация: В статье представлена методология, использующая Detectron2, систему обнаружения объектов, для распознавания и локализации компонентов и элементов в зоне функционирования промышленных роботов. В архитектуре используется Faster R-CNN - тип сверточной нейронной сети, предназначенной для быстрого и точного обнаружения объектов. На примере изображений протестирован предлагаемый алгоритм.

Ключевые слова: Искусственный интеллект, нейронные сети, обнаружение объектов, промышленная робототехника, робототехника, машинное обучение.

Введение. В эпоху автоматизации и цифровой трансформации роль роботизированных систем в производстве как никогда высока. Поскольку промышленные предприятия во всем мире стремятся к повышению эффективности и точности, внедрение роботизированных устройств в различные производственные процессы становится краеугольным камнем современного производства. Однако с интеграцией этих роботизированных систем возникает актуальная задача: отслеживание их перемещений и действий в режиме реального времени для оптимизации рабочих процессов и обеспечения безопасности. Традиционные методы отслеживания, такие как системы WIFI и GPS, хотя и эффективны во многих случаях, но оказываются недостаточными в специализированных промышленных условиях, особенно в условиях ограничения связи, например, на предприятиях оборонной промышленности. Кроме того, динамика движения роботов в сложных условиях требует более сложного решения. В данной статье мы попытаемся исследовать тонкости использования систем технического зрения на основе камер в сочетании со вспомогательными датчиками для создания комплексного, эффективного и экономичного решения по отслеживанию роботов, подходящего для современных производственных условий. Таким образом, данная статья направлена на то, чтобы рассмотреть существующие возможности и перспективы преодолеть разрыв между потенциалом автоматизации роботов и проблемами, возникающими при их мониторинге в реальном времени.

Перспективы глубокого обучения, особенно в области обработки и расшифровки сложных визуальных данных, позволяют найти подходящее решение. В данном исследовании используются возможности Detectron2, передовой системы обнаружения объектов, для помощи в инициализации промышленных движущихся роботов.

1. Методология: генерация и пополнение данных для нейронных сетей

1.1. Сбор данных с помощью камер

Для создания эффективной системы машинного зрения необходимо иметь богатый и разнообразный набор данных. Для этого рекомендуют использовать несколько видеозаписей роботизированного аппарата в действии. Выбор ракурса сверху вниз, отражающего положение камеры наблюдения, обеспечивает всесторонний обзор и наиболее информативно фиксирует положение робота.



1.2. Извлечение и аннотирование изображений.

Из этих видеоданных извлекают отдельные кадры, превращая их в статичные изображения. Каждое изображение сопоставляется с файлом метаданных. В этот файл заносятся такие важные сведения, как тип объекта (в нашем случае - роботизированного устройства), относительные координаты центра робота и его относительные размеры.

1.3. Расширение набора данных.

При огромных объемах данных, необходимых для обучения надежной нейронной сети, ручной сбор данных в таких масштабах становится нецелесообразным. Здесь применяют методы дополнения данных. Применяются такие преобразования, как:

- горизонтальное и вертикальное отражение,
- размытие изображения,
- введение шумов,
- масштабирование.

Нередко применяются методы синтетического расширения набора данных. Это не только обогащает набор данных, но и обеспечивает инвариантность нейронной сети к подобным преобразованиям, повышая ее устойчивость.

1.4. Разбиение массива данных.

После того как расширенный набор данных готов, очень важно разумно разделить его для эффективного обучения и проверки нейронной модели. Как правило, 70% данных отводится для обучения, где происходит собственно обучение. 20% отводится для тестирования, что позволяет убедиться в том, что обучение модели идет правильно. Оставшиеся 10% служат валидационным набором, выполняющим функцию окончательной проверки и оценивающим работу модели на совершенно неизвестных данных.

2. Разработка алгоритма и настройка модели

В исследовании используется библиотека Detectron2, включающая в себя набор моделей, оптимизированных для задач обнаружения объектов. Выбор Faster R-CNN с основой ResNet-101 обусловлен тем, что она доказала свою точность и вычислительную эффективность.

Выбор фреймворка Detectron2 обусловлен тем, что он поддерживает различные современные архитектуры обнаружения объектов. Среди них используется Faster R-CNN с основой ResNet-101.

Архитектура Faster R-CNN может быть представлена в виде:

$$Y = f(X; \Theta),$$

где: Y - выходной сигнал, содержащий обнаруженные ограничительные рамки и оценки классов,

X - входное изображение,

Θ - параметры модели.

Модель настраивается на работу либо с графическим процессором (GPU), либо с центральным процессором (CPU) в зависимости от доступности, что обеспечивает эффективность вычислений.

Для отсеивания неопределенных обнаружений задается порог достоверности. Для этого используется пороговое значение τ , равное 0,5, то есть достоверными считаются только те обнаружения, доверительный балл которых превышает 0,5.

Ускоренная R-CNN, как следует из названия, оптимизирует обычную R-CNN путем введения сети предложения регионов (RPN), которая совместно с сетью обнаружения использует сверточные функции полного изображения, что позволяет практически без затрат предлагать регионы.



Математически RPN работает путем скольжения небольшого окна по конволюционной карте признаков входного изображения. В каждом месте скользящего окна предсказывается несколько предложений регионов, каждому из которых присваивается балл объектоориентированности, определяемый следующим образом:

$$S = \sigma (W \cdot f + b),$$

где: S - оценка объектоориентированности,

f - признак в скользящем окне,

W - выученные веса,

b - смещение,

σ - сигмоидная функция активации.

Обнаруженные объекты накладываются на исходное изображение с помощью ограничивающих рамок. Эти рамки параметризуются координатами центра, шириной и высотой:

$$B = (x_c, y_c, w, h)$$

где (x_c, y_c) - координаты центра, а w и

h - ширина и высота, соответственно.

Основу набора данных составляет разнообразный набор изображений из различных сред. Эти изображения, представленные в различных форматах, таких как .jpg, .jpeg, .png и .webp, подчеркивают адаптивность нашего метода.

Ошибки, если они возникают, регистрируются, не вызывая сбоев в работе системы, что обеспечивает непрерывность работы.

Алгоритм и сама модель реализованы на Python, псевдокод алгоритма представлен ниже:

- 1) Загрузка данных.
- 2) Конфигурация Detectron2.

Инициализируются конфигурации Detectron2 по умолчанию (cfg).

3) Установка вычислительного устройства. В качестве вычислительного устройства GPU (если доступно) или CPU.

4) Объединение конфигурации Faster R-CNN с конфигурацией ResNet-101 из парка моделей Detectron2.

5) Установка порога доверия для обнаружения объектов.

6) Загрузка в конфигурацию предварительно обученные веса Faster R-CNN with ResNet-101 из парка моделей.

7) Инициализация модели.

Инстанцируется предиктор (предсказатель), с использованием указанной конфигурации.

8) Обнаружение и визуализация объектов.

9) Обработка ошибок. Если при обработке изображения возникла ошибка (например, неверный формат файла, поврежденное изображение), перехватить исключение и вывести сообщение об ошибке вместе с путем к проблемному изображению.

Результаты моделирования.

Модель была протестирована на введении изображений с несколькими роботами в закрытой среде, чтобы оценить ее работу в загроможденной обстановке. Модель эффективно локализовала роботов без посторонних аннотаций, как видно на изображении 1.



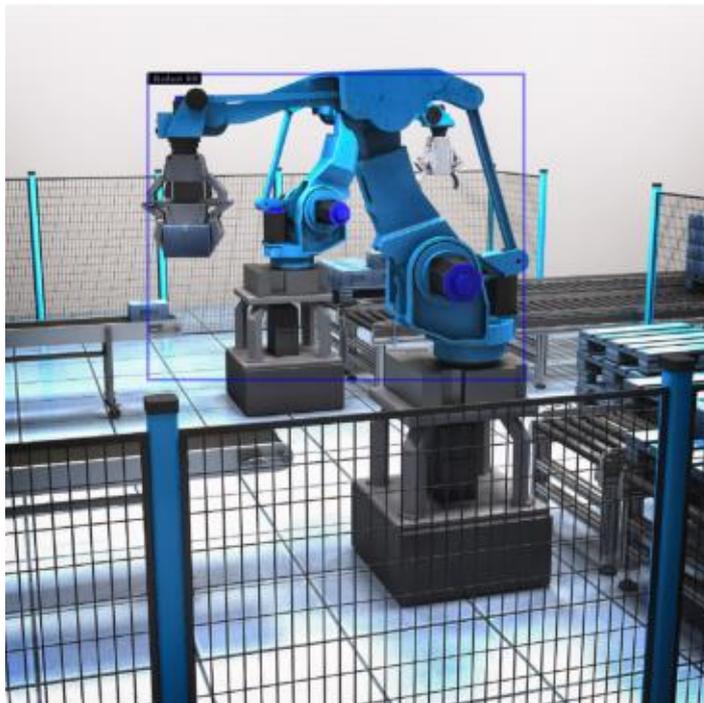


Рисунок 1 – Пример индикации робота на выборке, где заметны два робота

Помимо простой идентификации, модель успешно определяет конкретные компоненты роботов, как это видно на примере обнаружения движущейся части на рисунке 2. Это свидетельствует о более глубоком понимании и может оказаться полезным в приложениях, требующих детального анализа роботов.

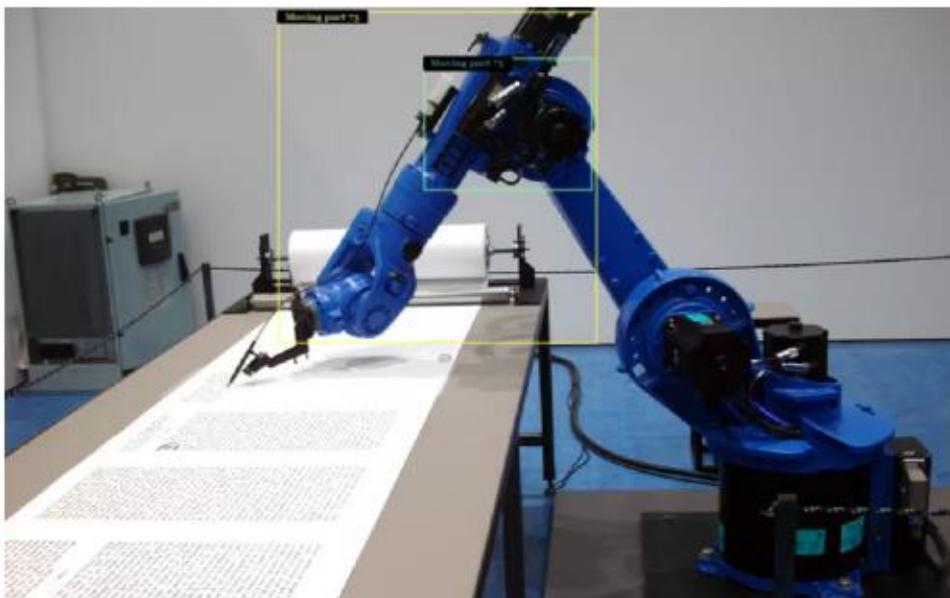


Рисунок 2 – Пример работы модели при сегментации движущихся частей робота

В целом модель показывает достаточную эффективность и минимальное количество ложных срабатываний и отрицательных результатов, что свидетельствует о надежности модели в плане точности обнаружения.

Заключение.

Проведенная оценка модели показывает достаточно хорошие возможности распознавания и классификации сложной техники, такой как промышленные роботы для представленного механизма использования описанного алгоритма. Благодаря достаточно точному построению ограничительных рамок и возможности идентификации конкретных компонентов (до 78 %) такие модели могут сыграть ключевую роль в развитии отраслей, основанных на робототехнике и автоматизации.

Техническая реализация такой модели предполагает обучение глубокой сверточной нейронной сети на обширном наборе данных изображений роботов, обеспечивающем разнообразие типов машин и условий эксплуатации. Для успешного развертывания такой системы в режиме реального времени на производстве необходимо обеспечить надежную аппаратную инфраструктуру, включая периферийные устройства, оснащенные специальными ускорителями интеллектуального кода и графическими процессорами. Это обеспечит оперативную обработку изображений и принятие решений на месте. Кроме того, для передачи данных, особенно если модель работает на нескольких предприятиях, необходима надежная сетевая инфраструктура. Интеграция с существующими производственными системами, такими как MES и SCADA, обеспечит бесперебойную работу, а регулярные обновления программного обеспечения позволят поддерживать точность и эффективность системы на должном уровне. В промышленных условиях такого рода модели могут помочь в обслуживании роботов, и обеспечении качества их функционирования и безопасности взаимодействия с операторами линии. Кроме того, такого рода модели позволяют обеспечить мониторинг контроля качества оценивая работу промышленных робототехнических средств в пределах заданных параметров.

Список литературы:

1. Wu, Y., Kirillov, A., Massa, F., Lo, W.Y., & Girshick, R. (2019). Detectron2.
2. Lin, T.Y. *et al.* (2014). Microsoft COCO: Common Objects in Context. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds) Computer Vision – ECCV 2014. ECCV 2014. Lecture Notes in Computer Science, vol 8693. Springer, Cham. https://doi.org/10.1007/978-3-319-10602-1_48

