**Анализ предметной области**

На любом производстве есть факторы, опасные для жизни или здоровья работников. Такие факторы могут быть физическими, химическими или биологическими. Предотвращение последствий их воздействия зависит от грамотного, надежного устройства, а также действенности системы ОТ на предприятии.

Охрана труда – это система различных мероприятий: юридических, социально-экономических, организационно-технических, санитарно-гигиенических, профилактических, проводимых с целью уберечь жизнь, сохранить здоровье работающего, повысить качество условий его трудовой деятельности, минимизировать воздействие негативных факторов производственной среды. В свою очередь, правила техники безопасности на предприятии – это комплекс организационных и технических мероприятий, имеющих целью предотвратить несчастные случаи, создать безопасные условия на местах выполнения рабочих обязанностей.

Соблюдение этих правил обязательно, однако не все сотрудники этого придерживаются, в результате чего на предприятиях могут происходить несчастные случаи, из-за которых страдают как сотрудники, получившие травмы, так и предприятия, вынужденные платить штрафы, искать замену работникам и т.д.

Одним из самых простых, но достаточно важных пунктов техники безопасности является «соблюдение правил ношения средств индивидуальной защиты (СИЗ)». Видов спецодежды насчитывается достаточно много, например: специальные защитные костюмы и обувь, жилеты, комбинезоны, каски, перчатки/рукавицы и многое другое. Набор СО обычно зависит от типа производства, самыми же часто используемыми являются каски и жилеты. В качестве примера работы модуля будем провертеть соблюдение ношения каски.

В приложении к Приказу Министерства здравоохранения и социального развития Российской Федерации от 9 декабря 2009 приведены перечни необходимой экипировки для работников на предприятиях нефтегазовой отрасли. Пример СИЗ представлен на рисунке 1.



Рисунок 1 – Пример ношения средств индивидуальной защиты (СИЗ)

Сегодня существует много подходов, позволяющих улучшить качество соблюдения норм безопасности на предприятиях. Например, чтобы отслеживать соблюдение правил ношения СИЗ можно использовать технологии компьютерного зрения.

**Выбор моделей машинного обучения**

Компьютерное зрение — это область компьютерных наук, которая стремится расширить возможности компьютеров по идентификации и определению объектов и людей на изображениях и видео. С его помощью мы можем определить положение человека на изображении, а также понять одеты ли на него все необходимые средства защиты.

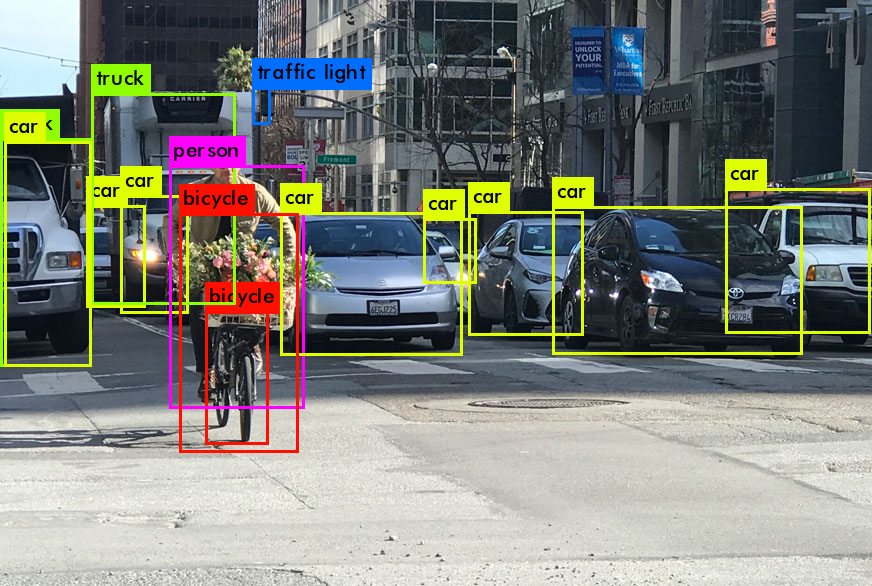


Рисунок 2 – Пример детектирования объектов на изображении

Свести описанную выше задачу можно к задаче детектирования объектов –задаче машинного обучения, в рамках которой выполняется определение наличия или отсутствия объекта определённого домена на изображении и нахождение границ этого объекта в системе координат пикселей исходного изображения. В зависимости от алгоритма обучения, объект может характеризоваться координатами ограничивающей рамки, ключевыми точками или контуром объекта.

Подходов к решению задачи детектирования объектов достаточно много. Их можно разделить на два типа: классические и на основе нейронных сетей. Так как первые проигрывают вторым по качеству работы, остановимся на нейронных сетях. Они тоже имеют классификацию, на сегодня существует два параллельно развивающихся направления:

* Двухэтапные методы (англ. two-stage methods) – подход, разделённый на два этапа. На первом этапе селективным поиском или с помощью специального слоя нейронной сети выделяются регионы интереса (англ. regions of interest, RoI) — области, с высокой вероятностью содержащие внутри себя необходимые объекты. На втором этапе выбранные регионы рассматриваются классификатором для определения принадлежности исходным классам и регрессором, уточняющим местоположение ограничивающих рамок.
* Одноэтапные методы (англ. one-stage methods) — подход, не использующий отдельный алгоритм для генерации регионов, вместо этого предсказывая координаты определённого количества ограничивающих рамок с различными характеристиками, такими, как результаты классификации и степень уверенности и в дальнейшем корректируя местоположение рамок.

В свою очередь одноэтапные методы выигрывают двухэтапные, так как предсказание регионов хоть и обеспечивает хорошую точность, может быть очень затратным по времени процессом. Самой же популярной и развивающейся одноэтапной архитектурой нейронных сетей для детектирования объектов на изображении является YOLO.

YOLO (“You Only Look Once”) — современный алгоритм глубокого обучения, который широко используется для обнаружения объектов. Он был разработан Джозефом Редмоном и Али Фархади в 2016 году. Основное отличие YOLO от других алгоритмов сверточных нейронных сетей (CNN), используемых для обнаружения объектов, заключается в том, что он очень быстро опознает объекты за счет того, что YOLO подразумевает ввод сразу всего изображения, и оно проходит через нейронную сеть только один раз. Именно поэтому он называется “Стоит только раз взглянуть”. В других алгоритмах этот процесс происходит многократно, то есть изображение проходит через CNN снова и снова. Таким образом, можно сделать вывод, что YOLO обладает преимуществом высокоскоростного обнаружения объектов, чем не могут похвастать другие алгоритмы.

Для проведения экспериментов с архитектурой мной была выбрана 5 версия – YOLOv5. Она имеет реализацию на PyTorch и встроена в одноименный модуль для Python3, который можно установить с помощью pypi. Этот модуль предоставляет очень хорошую инфраструктуру как для обучения модели, так и для тестирования с построением графиков всех ключевых показателей. Перед тем как приступить к обучению и экспериментам нам необходимо найти, собрать и подготовить данные.

**Выбор наборов данных для обучения**

Для обучения детектора нам необходимо собрать и подготовить данные. Прежде всего определим классы объектов, местоположение которых мы хотим найти. В первую очередь нам необходимо найти человека на кадре – он будет первым классом. Далее определим классы для каждого вида СИЗ. На данном этапе разработки модуля остановимся на детектировании защитной каски. Такой выбор обусловлен несколькими факторами: Во-первых, данный класс будет присутствовать в большинстве вариантов использования модуля, так как каски часто используются на промышленных предприятиях. Во-вторых, детектирование мелких элементов экипировки является более сложной задачей. Ракурс камер на предприятиях обычно не позволяет разглядеть их наличие даже человеку, соответственно требуется дополнительная настройка оборудования и взаимодействие с заказчиком, для определения дополнительных классов и получения данных для обучения или разрешения на их сбор и обработку.

Для обучения детектора людей мной был выбран датасет CrowdHuman. Он представлен в открытом доступе и содержит 15000 фотографий, на которых представлено ~340 тысяч объектов класса «человек». Ссылка для ознакомления и скачивания: <https://www.crowdhuman.org/>.

Для обучения детектора касок я произвел сбор и разметку данных из открытых источников. В процессе поиска мой выбор пал на завод металлоконструкций Аполло. Компания предоставляет потоки видео с IP камер, расположенных на их предприятии, в открытом доступе. Посмотреть их можно на сайте: <https://apollo-zmk.ru/about/interesnoe/live-video/>. Мной были записаны небольшие видео фрагменты с нескольких камер, из которых в последствии получилось собрать и разметить ~500 фотографий. После чего я применил техники аугментирования данных и получил набор, состоящий из ~1500 фотографий.

**Обучение моделей и проведение экспериментов**

На первом этапа создадим детектор людей. Для решения этой задачи мной были выбраны 3 архитектуры YOLOv5 – модели nano, small и medium. Ниже представлена таблица, содержащая метрики, полученные в результате обучения. Мной были выбраны:

* Precision/точность – отношение количества истинно положительных предсказаний ко всем положительным предсказаниям. Показывает какая доля объектов, названных классификатором положительными действительно положительная.
* Recall/полнота – отношение количества истинно положительных предсказаний к сумме истинно положительных и ложно отрицательных. Показывает какую долю объектов положительного класса из всех объектов положительного класса нашел алгоритм.
* mAP – средняя точность, вычисляемая для всех классов. 0.5 – пороговое значение для IoU (Intersection over Union) – показывает процент пересечения двух областей. Если пересечение больше 0.5 – то считаем, что объект найден правильно. 0.5:0.95 – означает средняя значение mAP с пороговыми значениями от 0.5 до 0.95

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Модель | Precision | Recall | mAP 0.5 | mAP 0.5:0.95 |
| YOLOv5\_nano |  |  |  |  |
| YOLOv5\_small |  |  |  |  |
| YOLOv5\_medium |  |  |  |  |

Источники

1. <https://www.trudohrana.ru/article/103436-qqq-17-ohrana-truda-i-tehnika-bezopasnosti-na-predpriyatii>
2. <https://neerc.ifmo.ru/wiki/index.php?title=Задача_нахождения_объектов_на_изображении>
3. <https://medium.com/nuances-of-programming/алгоритм-yolo-простым-языком-a4e666be4d41>