МЕТОД ОБНАРУЖЕНИЯ И РАСПОЗНАВАНИЯ ДВИЖУЩИХСЯ ОБЪЕКТОВ В ВИДЕОПОТОКЕ

Альховский А.А., Деменковец Д.В.

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Серебряная Л.В., доцент, к.т.н.

В работе представлен метод решения задачи обнаружения и распознавания движения объектов для обеспечения требуемого уровня безопасности на промышленных производствах. Использованы глубокая искусственная нейронная сеть, одноплатный компьютер, алгоритмы обнаружения движения и сопровождения объектов.

Производственный процесс, в котором участвует человек, несет в себе опасность получения травм. Соблюдения одних лишь требований безопасности недостаточно. Во всем мире ежегодно регистрируется примерно 340 млн. несчастных случаев на производствах. Своевременное обнаружение нестандартных ситуаций может помочь спасти жизнь человека.

В настоящее время является актуальным автоматизированный контроль процесса производства и объектов, участвующих в нем, включая людей. Были определены нормальные режимы производственного процесса, характеризующие безопасное состояние человека в нем.

В работе представлен метод обнаружения движения объектов в видеопотоке с применением одноплатного компьютера Jetson NANO [1] и искусственных нейронных сетей (ИНС). Для достижения поставленной цели необходимо решить следующие задачи: определение состояния объекта и направление его движения. Анализ существующих способов показал, что решение сформулированной задачи на основе датчиков движения, установленных в касках или на оборудовании, достаточно дорогое. Поэтому было решено использовать модель на основе ИНС, данные, получаемые с видеокамеры, установленной на одноплатном компьютере Jetson NANO [1]. Jetson NANO позволяет работать в режиме, близком к режиму реального времени, и имеет графическую мощность, равную 128 ядер, достаточной для работы с глубокой ИНС.

Решение задачи начинается с определения состояния объекта, т.е. движется он или нет. В случае движения его необходимо распознать в видеопотоке, а также уметь отличить от других объектов, присутствующих в видеопотоке. Затем требуется определить, изменялось ли местоположение объекта в видеопотоке, полученном со статической камеры. В качестве обучающей выборки для ИНС использован набор данных под названием COCO (Common Objects in Context) [2]. Набор данных COCO содержит ~64000 размеченных изображений с людьми и отличается простым форматом данных. Для обнаружения объекта в кадре был использован алгоритм YOLO (You Only Look Once) [3]. Этот алгоритм требует получение из исходного кадра признаков изображений. Преобразование исходного кадра в многомерную матрицу признаков выполняется на основе архитектуры ИНС MobileNet v2 [4]. Результат работы алгоритма YOLO представляет собой многомерную матрицу предсказаний ИНС, которая носит избыточный характер для отображения. Поэтому необходимо использовать преобразование выходов ИНС. Результаты оцениваются вероятностями и в дальнейшем не учитываются те из них, которые имеют вероятность меньше 0.5. Для сокращения избыточности оставшихся выходных значений использовался дополнительный критерий отбора.

Результат полученного решения – 2-3 кадра в секунду при обнаружении движущегося объекта. Данное решение является недостаточно производительным для безопасности производства. Чтобы повысить производительность обнаружение выполняется один раз в 50 кадров. В промежутке между обнаружениями используется алгоритм сопровождения объектов в видеопотоке на основе многомерной корреляционной фильтрации [5]. Использование алгоритма сопровождения объектов по видеопотоку позволило обновлять местоположение объекта для каждого кадра, затрачивая меньше вычислительной мощности. Использование значения координат центра, ширины, высоты позволило реализовать алгоритм идентификации объекта, основанный на координатах центра объекта. Это позволило распознать объект в видеопотоке.

Использование алгоритма сопровождения объектов в видеопотоке позволило обновлять местоположение объекта для каждого кадра, затрачивая меньше вычислительной мощности. Использование вышеописанного метода позволило достичь скорости в 12 кадров /секунда.

Список использованных источников:

1. Jetson Nano Developer Kit // NVIDIA AUTONOMOUS MACHINES [Электронный ресурс]. – Режим доступа: https://developer.nvidia.com/embedded/jetson-nano-developer-kit. – Дата доступа: 20.03.2020.
2. Microsoft COCO: Common Objects in Context // Cornell University [Электронный ресурс]. – Режим доступа: https://arxiv.org/abs/1405.0312/ Дата доступа: 20.03.2020
3. YOLOv3: An Incremental Improvement // Cornell University [Электронный ресурс]. – Режим доступа: https://arxiv.org/abs/1804.02767/ Дата доступа: 20.03.2020
4. Accurate Scale Estimation for Robust Visual Tracking / Martin Danelljan [и др.]. BMVA Press, 2014.