

## МОДУЛЬ СИТУАЦИОННОЙ ВИДЕОАНАЛИТИКИ ТОРГОВОГО ЗАЛА

**Аннотация.** В статье представлена разработка модуля ситуационной видеоаналитики, реализующего выявление «одинокого покупателя» в торговом зале магазина компании «Строительный двор», а также отправку соответствующего сообщения в корпоративный Telegram-чат сотрудников. Реализация такого функционала будет способствовать повышению качества обслуживания покупателей и росту показателей эффективности торговой точки.

**Ключевые слова:** видеоаналитика, компьютерное зрение, искусственные нейронные сети, детектирование объектов, сценарий «одинокий покупатель».

### Основные определения

Ситуационная видеоаналитика — видеоаналитика, предназначенная для анализа ситуаций и (или) сценариев в сцене видеонаблюдения.

Ситуация — соответствие наблюдаемой на сцене видеонаблюдения совокупности (количественных и качественных) изменений или их отсутствия заданному описанию, подлежащее обнаружению системой видеоаналитики.

Сценарий ситуации — заданная совокупность или последовательность взаимосвязанных событий в сцене видеонаблюдения, характеризующая ситуацию.

Сцена видеонаблюдения — пространство в поле зрения видеокамеры.

**Введение.** Успешное ведение любого бизнеса связано с принятием обоснованных управленческих решений. Правильность принимаемых решений оценивается на основании показателей эффективности, характеризующих данный вид деятельности. Настоящее исследование выполнялось в контексте решения задачи управления торговым залом магазина компании «Строительный двор». Одним из основных показателей эффективности торговой точки является коэффициент конверсии.

Конверсия — это важное (целевое) действие, которое совершает посетитель магазина. Например: совершение покупки, оформление подписки, добавление товара в корзину (в интернет-магазине), регистрация. Коэффициент конверсии  $CR$  — это отношение количества  $N_c$  посетителей, выполнивших действие, к общему количеству посетителей  $N$ , выраженное в процентах [1]:

$$CR = \frac{N_c}{N} \cdot 100\% \quad (1)$$

Увеличение данного показателя, в свою очередь, обеспечивает бизнесу рост других показателей эффективности: увеличение прибыли, ускорение возврата инвестиций, и др. Повышению коэффициента конверсии могут способствовать различные элементы мерчандайзинга (установка подходящего торгового оборудования, оформление торгового зала и т. п.). При этом следует учитывать, что коэффициент конверсии связан с понятием лояльности со стороны потребителей. Ключевым фактором, влияющим на лояльность клиентов, является степень их удовлетворенности от предыдущих обращений в компанию. Удовлетворенность клиентов — лучший показатель для определения вероятности совершения потребителем повторной покупки. Более того, удовлетворенный покупатель, скорее всего, будет рекомендовать

продукцию компании своим знакомым, друзьям и родственникам, что является мощнейшим маркетинговым преимуществом [2].

Повышению степени удовлетворенности клиентов и уровня конверсии способствует работа консультантов магазина с клиентами. Успешное консультирование позволяет не только продать товар, но и создать положительное впечатление о магазине и привлечь потенциальных клиентов в дальнейшем.

Для контроля и повышения качества консультирования могут применяться современные технологии с использованием видеоаналитики. В частности, в настоящем исследовании рассматривается реализация сценария «одинокий покупатель». Она включает в себя определение ситуаций, при которых покупатель значительное время находится без внимания со стороны консультанта. Определение таких ситуаций выполняется на основе анализа видеопотока с камер, установленных в торговом зале.

**Проблематика исследования.** В рамках данной работы была поставлена цель повысить качество обслуживания покупателей магазинов компании «Строительный двор» путем разработки и внедрения модуля ситуационной видеоаналитики. Качество обслуживания в контексте данного исследования характеризуется временем, в течение которого посетитель торгового зала ожидает сотрудника — консультанта магазина. Разрабатываемый модуль должен обеспечивать выявление «одинокого покупателя» — посетителя торгового зала, рядом с которым в течение заданного промежутка времени не фиксируется появление консультанта магазина, и автоматическую отправку сообщения об этом в корпоративный Telegram-чат сотрудников торговой точки.

В ходе разработки модуля требовалось решение ряда задач:

- реализации «нарезки» видеоряда, полученного с камеры торгового зала, на отдельные кадры;
- реализации системы детектирования объектов — покупателей и сотрудников магазина на видеокадрах:
  - создания обучающего набора изображений путем фильтрации и ручной разметки множества видеокадров;
  - обучения искусственной нейронной сети, выполняющей детектирование объектов (людей) на видеокадрах;
  - реализации функции реидентификации — повторной идентификации объекта на последовательности видеокадров как одного и того же объекта с учетом определения нового положения объекта в кадре при его перемещении или смене области зрения видеокамеры;
  - обучения искусственной нейронной сети, выполняющей классификацию выявленных объектов (разделение людей на 2 класса: покупателей и сотрудников магазина);
- реализации алгоритма определения наличия контакта (посетителя и консультанта) в течение установленного промежутка времени;
- реализации автоматической отправки сообщения в Telegram-чат.

Изучение работ по созданию систем детектирования и классификации объектов на изображениях [в частности, 3-6] позволило определить основные подходы к решению задач

анализа изображений, а также определить подобрать архитектуры нейронных сетей, наиболее подходящих для целей исследования.

**Материалы и методы.** В процессе разработки модуля использовались видеоматериалы, полученные с камер видеонаблюдения сети магазинов «Строительный Двор». Видеопоток состоит из серии последовательных изображений (кадров). Следует отметить, что камеры, установленные в магазинах компании, не имеют встроенного функционала детектирования объектов на изображениях, поэтому требовалась разработка соответствующего инструментария.

#### *Извлечение исходных изображений*

Извлечение кадров с видеокamеры реализовано с помощью библиотеки OpenCV [7]. Для обеспечения возможности работы модуля видеоаналитики в режиме реального времени реализована передача кадров с заданным интервалом. Несмотря на то, что для процесса реидентификации объектов использовалась одна из самых быстрых моделей (SSD), данный процесс выполняется достаточно долго, и технические возможности системы не позволяют обрабатывать 24 кадра в секунду. Опытным путем было установлено, что при ограничении на обработку 6 кадров в секунду объекты не успевают изменить свою позицию настолько, чтобы ухудшились показатели работы алгоритма реидентификации. В то же время, такая скорость передачи данных позволяет выполнять обработку в режиме реального времени.

Таким образом, было принято решение использовать не все кадры видеопотока, а каждый четвертый кадр. Был разработан алгоритм для «нарезки» видеоряда на кадры. На вход алгоритма подается путь к видеозаписи и значение параметра «необходимое количество кадров в секунду»; на выходе — архив с нарезанными кадрами.

#### *Детектирование посетителей и сотрудников*

Задача обнаружения на видеокadрах посетителей и сотрудников магазина решалась в два этапа:

- 1) обнаружение (детектирование) людей без определения их принадлежности к тому или иному классу;
- 2) бинарная классификация объектов, обнаруженных (детектированных) на первом этапе: отнесение каждого человека к классу посетителей либо консультантов магазина.

Для реализации каждого этапа использовалась своя нейронная сеть.

Детектирование объектов — задача, в рамках которой необходимо выделить несколько объектов на изображении посредством нахождения координат рамок, ограничивающих эти объекты, а также классификации ограничивающих рамок на множестве заранее известных классов. Применение отдельной нейронной сети для детектирования (решения задачи первого этапа) было необходимо для наиболее четкого определения ограничивающих рамок объектов с целью облегчения их будущей реидентификации.

По результатам сравнительного анализа различных подходов к реализации детектирования людей на кадрах видеоряда был выбран подход, основанный на использовании предобученной глубокой нейронной сети семейства архитектур YOLO. При выборе архитектуры сети учитывалась не только точность результатов детекции, но и необходимость высокой скорости обнаружения объектов при работе в режиме реального времени.

Из хранилища моделей Pytorch [8], была загружена модель yolov5s от Ultralytics. Эта модель предобучена на данных из открытого источника COCO (common objects in context) [9]

и осуществляет поиск базовых классов (люди, животные, автомобили и др.). На вход нейронной сети подается изображение, на выходе — классы объектов и рамки, ограничивающие объекты разных классов. При использовании этой сети на первом этапе решения (до применения алгоритма классификации на втором этапе) каждый объект (человек) считался неопределенным (неклассифицированным).

Классификация также выполнялась с помощью сети yolov5s от Ultralytics, но для определения класса объекта — «посетитель»/«сотрудник» — было реализовано обучение сети на специально подготовленном обучающем наборе изображений. На вход классифицирующей сети подавались изображения и рамки, ограничивающие объекты, на выход сеть должна передать метки классов объектов и степень уверенности в ответе. В случае, если степень уверенности составила более, чем 0.5, информация об объекте (класс и степень уверенности) сохранялась в базе данных.

На этапе классификации было введено ограничение на минимальный размер объекта. Это связано с тем, что при попадании в кадр слишком малой части объекта нельзя достоверно определить класс этого объекта (посетитель или сотрудник).

#### ***Создание обучающей выборки для классификатора***

Создание обучающей выборки производилось на основе «нарезанных» кадров исходного видеопотока. Основные этапы следующие.

- Ручная фильтрация изображений по признаку наличия людей в сцене камеры. Были выбраны и сохранены неповторяющиеся кадры.
- Разметка отфильтрованных изображений. Для разметки использовался инструментальный платформ Roboflow [10].
- Ручное создание ограничивающих рамок для каждого объекта на каждом кадре с указанием класса объекта.

Таким образом был сформирован набор данных, состоящий из 1000 обработанных кадров и аннотаций к ним. Набор данных был разделен на обучающую, валидационную и тестовую выборки в соотношении 70%, 20% и 10% соответственно. С помощью Roboflow был также сгенерирован код и api-ключ для дальнейшего доступа к набору данных через среду Colab.

#### ***Обучение классификатора***

Обучение нейронной сети-классификатора проходило в течение 100 эпох. Для определения качества идентификации объектов использовалась метрика Intersection over Union (IoU) [11]. Для двух непустых множеств  $A$  и  $B$ , значение IoU определяется как

$$\text{IoU}(A, B) = \frac{|A \cap B|}{|A \cup B|}, \quad (2)$$

где  $A$  — ограничивающая рамка, определенная нейронной сетью, а  $B$  — ограничивающая рамка, определенная разметкой.

По итогам обучения были сохранены лучшие и последние веса. Для классификации использовалась нейронная сеть с весами, показавшими лучший результат. На рис. 1 и 2 представлены примеры результатов работы обученной сети-классификатора.

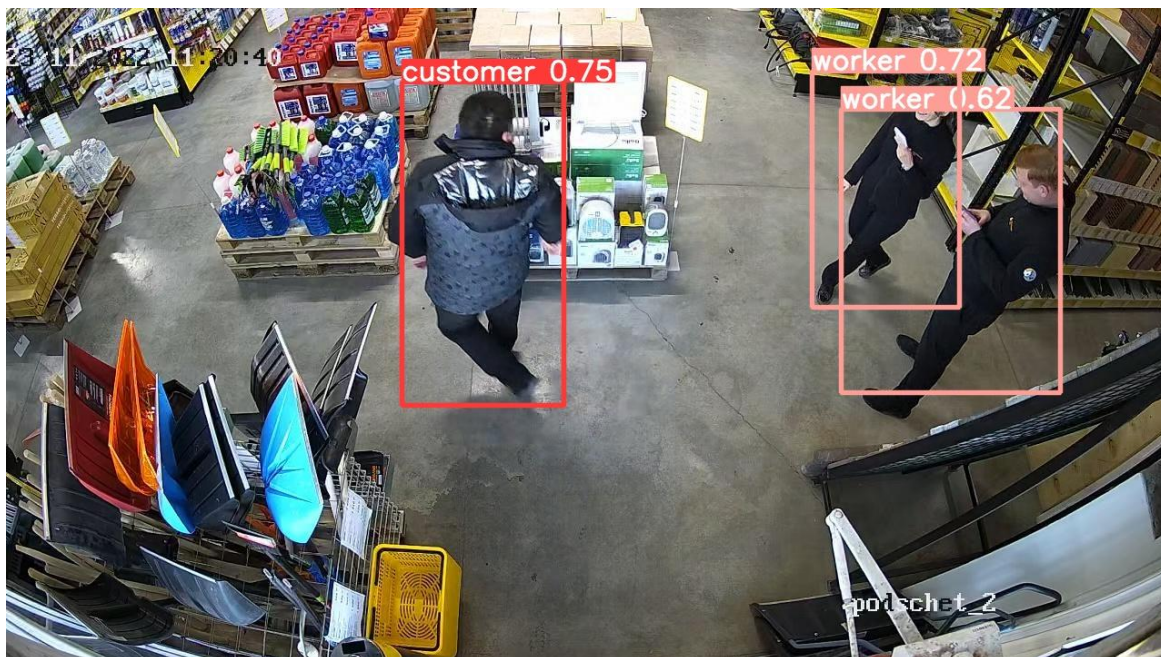


Рис. 1. Пример работы классифицирующей нейронной сети

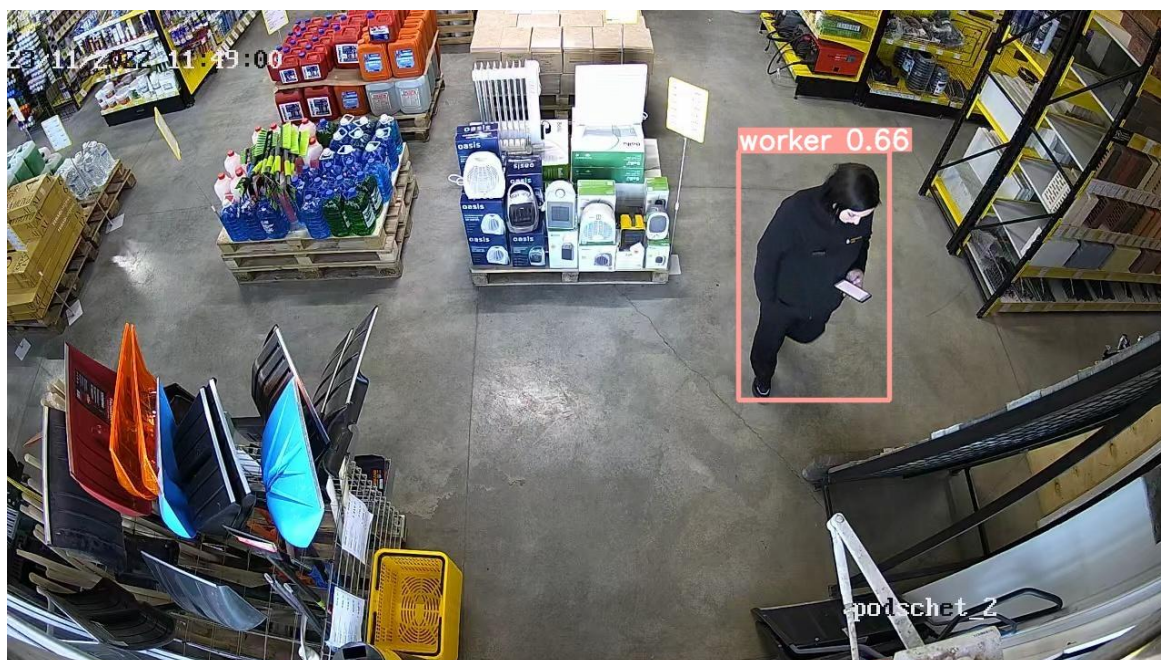


Рис. 2. Пример работы классифицирующей нейронной сети

### ***Алгоритм определения наличия контакта***

Для определения наличия контакта между объектами в ситуационной видеоаналитике используются алгоритмы анализа движения объектов (анализируют перемещения объектов и расстояние между ними) или алгоритмы компьютерного зрения для обнаружения приближений и контакта объектов.

С учетом имеющихся технических возможностей в данном исследовании было принято решение: для определения контакта между посетителем и консультантом не использовать



третью нейронную сеть, а реализовать алгоритм, определяющий расстояние между людьми. Алгоритм анализирует следующие параметры: классы присутствующих на изображении людей (должны присутствовать представители обоих классов), расстояние между объектами и продолжительность нахождения объектов на близком расстоянии.

Проверка осуществляется следующим образом: клиент считается проконсультированным, если консультант находился в непосредственной близости к покупателю на протяжении не менее 4 секунд. В противном случае (консультант просто проходил мимо) факт консультации зафиксирован не будет.

Расстоянием между консультантом и покупателем считается расстояние между серединами рамок, ограничивающих данные объекты на изображении. Расстояние считается близким, если между серединами рамок не более 250 пикселей. На рис. 3 показан пример определения расстояния.

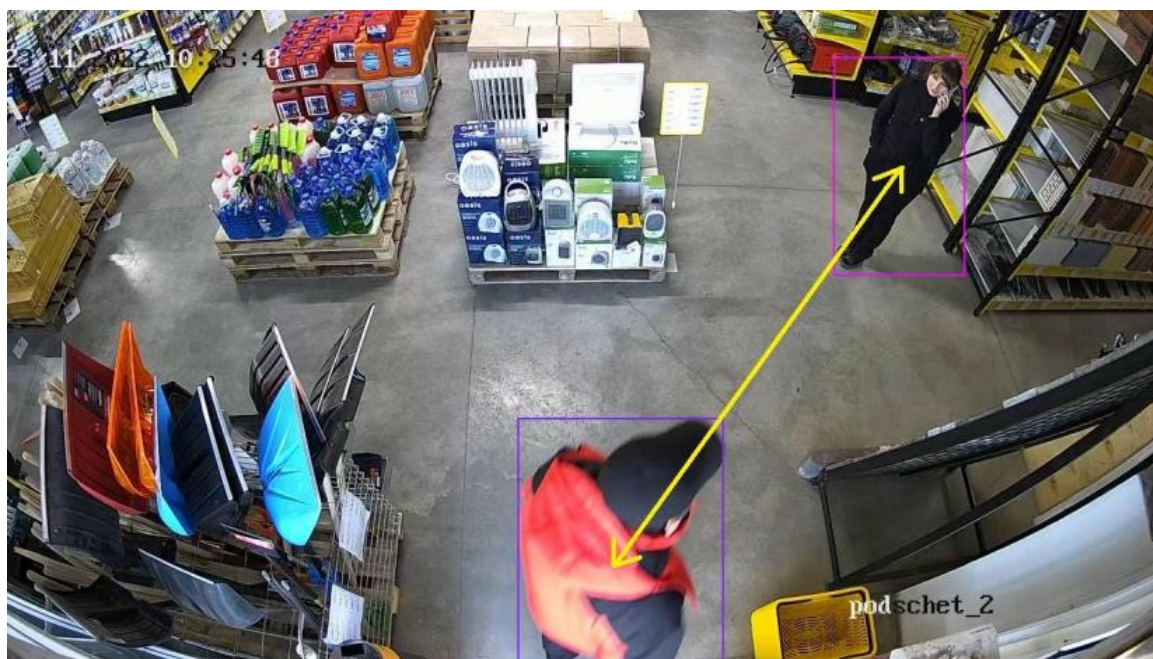


Рис. 3. Пример определения расстояния

### ***Реидентификация объектов***

Для реидентификации использовался алгоритм DeepSORT [3], реализованный в фреймворке `deep_sort_realtime` [12]. Алгоритм получает на вход изображение и рамки, ограничивающие объекты на этом изображении (получены с помощью нейронной сети, выполняющей детектирование объектов). Алгоритм определяет объект «сквозь кадры», сравнивая его шаблон с шаблонами других объектов.

### **Результаты**

#### ***Общий алгоритм работы модуля видеоаналитики***

- Получение видеопотока с камеры посредством `rtp`-соединения. Получение и обработка кадров видеопотока с помощью `OpenCV`.
- Считывание присланного с камеры кадра. В соответствии с установленным количеством обрабатываемых кадров в секунду — пропуск кадра либо начало его обработки.

- Передача изображения на вход детектирующей нейронной сети с установленными параметрами ширины и высоты изображения; минимальная степень уверенности, что детектирован человек, установлена равной 0,5.
- Реидентификация человека с использованием алгоритма DeepSORT (с параметрами  $\text{max\_cosine\_distance} = 0.3$  и  $\text{max\_age} = 5$ ).
- Если обнаруженный человек — новый в сцене, то занесение информации о нем и о новом посещении в базу данных.
- Проверка, определен ли класс у каждого из присутствующих в сцене видеонаблюдения.
- Если есть неопределенные объекты (с неизвестным классом), то оценка размеров таких объектов. Если размер объекта более 50 пикселей в высоту, то передача изображения с координатами объекта на вход классифицирующей нейронной сети. На выходе сети — класс объекта и степень уверенности. Если степень уверенности более 0.5, то сохранение обновленной информации об объекте (класс и степень уверенности) в базе данных.
- Обновление списка присутствующих в сцене видеонаблюдения и обновление количества кадров для посещений.
- Удаление из числа активных тех посещений, для которых id покупателей в течение 5 кадров не находятся в сцене видеонаблюдения.
- Проверка на наличие консультирования для каждого клиента с каждым из консультантов.
- Проверка на наличие «одинокоего покупателя»: если у покупателя нет зафиксированного контакта с консультантом в течение 30 секунд, то передача сообщения о наличии в торговом зале непроконсультированного покупателя в групповой Telegram-чат консультантов (посредством реализованного telegram-бота).
- Если в кадре присутствует новый человек (информация о котором еще не сохранена в базе данных), то обработка кадра и передача его на вход нейронной сети-классификатора для определения класса объекта и сохранение результатов классификации.
- Получение нового кадра.

**Заключение.** Разработан модуль видеоаналитики, реализующий функцию контроля за осуществлением консультирования посетителей торгового зала сотрудниками магазинов компании «Строительный двор». Предполагается, что такой контроль за действиями сотрудников будет способствовать повышению качества обслуживания покупателей и приведет к росту лояльности потребителей и повышению коэффициента конверсии.

## СПИСОК ЛИТЕРАТУРЫ

1. Показатели торговли. — Текст: электронный // ФИНОКО: Бюджетирование и управленческий учет: официальный сайт. — 2023. — URL: <https://www.finoko.ru/pokazateli-torgovli/> (дата обращения 28.05.2023).
2. Васин Ю.В. Эффективные программы лояльности. Как привлечь и удержать клиентов / Ю.В. Васин, Л.Г. Лаврентьев, А.В. Самсонов — Москва: Альпина Бизнес Букс, 2011. — 288 с. — Текст: непосредственный.
3. Wojke N. Simple Online and Realtime Tracking with a Deep Association Metric / N. Wojke, A. Bewley, D. Paulus. — URL: <https://arxiv.org/abs/1703.07402> (date of the application 28.05.2023). — Text: electronic.

4. Хурсов П.С. Алгоритмы детекции объектов для анализа изображений / П.С. Хурсов, Н.А. Искра. — URL: [https://libeldoc.bsuir.by/bitstream/123456789/37064/1/Khursov\\_Algoritmy.pdf](https://libeldoc.bsuir.by/bitstream/123456789/37064/1/Khursov_Algoritmy.pdf) (дата обращения 28.05.2023). — Текст: электронный.
5. Береснев А.П. Исследование сверточных нейронных сетей класса YOLO для мобильных систем детектирования объектов на изображениях / А.П. Береснев, И.В. Зоев, Н.Г. Марков. — URL: <https://www.graphicon.ru/html/2018/papers/196-199.pdf> (дата обращения 28.05.2023). — Текст: электронный.
6. Girshick R. Rich feature hierarchies for accurate object detection and semantic segmentation / R. Girshick, J. Donahue, T. Darrell, J. Malik. — URL: <https://arxiv.org/abs/1311.2524> (date of the application 28.05.2023) — Text: electronic.
7. OpenCV: открытая библиотека для работы с алгоритмами компьютерного зрения: [сайт]. — URL: <https://opencv.org/> (дата обращения 28.05.2023). — Текст: электронный.
8. YOLOV5. — Text: electronic // PyTorch: official website — 2023. — URL: [https://pytorch.org/hub/ultralytics\\_yolov5/](https://pytorch.org/hub/ultralytics_yolov5/) (date of the application 28.05.2023).
9. COCO: A large-scale object detection, segmentation, and captioning dataset: [сайт]. — URL: <https://cocodataset.org/#home> (date of the application 28.05.2023). — Text: electronic.
10. Roboflow: платформа для работы с алгоритмами компьютерного зрения: [сайт]. — URL: <https://roboflow.com/> (дата обращения 28.05.2023). — Текст: электронный.
11. Intersection over Union (IoU) for object detection. — Text: electronic // PyImageSearch: official website. — 2022. — URL: <https://pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/> (date of the application 28.05.2023).
12. deep-sort-realtime 1.3.2: открытая библиотека для работы с алгоритмами компьютерного зрения: [сайт]. — URL: <https://pypi.org/project/deep-sort-realtime/> (дата обращения 28.05.2023). Текст: электронный.