Общество с ограниченной ответственностью «Цифровые решения»

УДК: 621.397

Регистрационный №АААА-А19-119042390070-0

Инв. №02-2019

УТВЕРЖДАЮ

Генеральный директор

Матвеев Д.В.

" " 201\_г.

**М.П.**

ОТЧЕТ

о выполнении НИОКР по теме:

"Разработка алгоритмов детектирования объектов на видеоданных с вертикального ракурса камеры. Разработка алгоритмов трекинга объектов на видеоданных с вертикального ракурса камеры. Обучение и тестирование алгоритмов трекинга и детектирования на основе сверточной нейронной сети."

(договор 2962ГС1/45326 от 01.04.2019)

(промежуточный)

|  |  |
| --- | --- |
| Научный руководитель работ | Хрящев В.В.  подпись, дата |

Ярославль, 2019

1. **Реферат**

Отчет 32 стр., 13 илл., 1 табл., 0 приложений, 30 источников

детектирование объектов, сопровождение объектов, машинное обучение, компьютерное зрение, нейронные сети

Цель работы на втором этапе выполнения НИОКР – разработка алгоритмов детектирования и трекинга игроков на площадке с вертикального ракурса камеры. Алгоритмы построены на основе сверточной нейронной сети, выполнены операции обучения и тестирования. К полученным результатам относятся:

* для обучения и тестирования системы детектирования и трекинга была создана база изображений размером 1000 и 100 изображений, соответственно, размер которых составлял 2048x1536 пикселей;
* разработан алгоритм детектирования изображений баскетболистов с вертикального ракурса видеокамеры;
* разработан алгоритм трекинга расположения баскетболистов данных с вертикального ракурса видеокамеры;
* проведено тестирование полученных алгоритмов на спортивной площадке Ярославского государственного университета.

Областями применения разработанных алгоритмов являются комплексы спортивной видеоаналитики, которые будут использованы спортивными клубами, многофункциональными спортивными комплексами.

Работы по данному этапу НИОКР были выполнены в полном объеме в соответствии с техническим заданием и календарным планом. Полученные результаты будут использованы на следующих этапах НИОКР и в коммерческой деятельности ООО «Цифровые решения».

**СПИСОК ИСПОЛНИТЕЛЕЙ**

|  |  |  |
| --- | --- | --- |
| **Должность** | **Подпись** | **ФИО, номер раздела** |
| Генеральный директор, ООО «Цифровые решения», к.т.н. |  | Матвеев Д.В. (подразделы 3.1 – 3.3) |
| Научный руководитель работ, к.т.н. |  | Хрящев В.В. (подразделы 3.1 – 3.2) |
| Программист ООО «Цифровые решения» |  | Казина Е.М. (подразделы 3.2 – 3.3) |
| Консультант ООО «Цифровые решения» |  | Федькина А.А. (подразделы 3.2 – 3.3) |

**СОДЕРЖАНИЕ**

|  |  |
| --- | --- |
| **1. Реферат** | 2 |
| **2. Введение** | 5 |
| **3. Основная часть** | 8 |
| 3.1. Разработка алгоритмов детектирования объектов на видеоданных с вертикального ракурса камеры. | 8 |
| 3.1.1. Получение изображений с камеры | 8 |
| 3.1.2. Обнаружение фона | 10 |
| 3.1.3. Сопоставление с шаблоном для детектирования игроков на баскетбольной площадке | 12 |
| 3.1.4. Обнаружение отдельного игрока | 14 |
| 3.1.5. Обнаружение нескольких игроков | 17 |
| 3.2. Разработка алгоритмов трекинга объектов на видеоданных с вертикального ракурса камеры | 19 |
| 3.2.1. Система трекинга | 19 |
| 3.2.2. Исправление искажений купольной камеры | 21 |
| 3.2.3. Трекинг игрока | 22 |
| 3.3. Обучение и тестирование алгоритмов трекинга и детектирования на основе сверточной нейронной сети | 26 |
| **4. Заключение** | 29 |
| **5. Список использованных источников** | 31 |

**2. ВВЕДЕНИЕ**

В настоящее время в спортивном сообществе возрастает потребность в отслеживании информации об игроках, которая могла быть полезной тренерам и ученым в области спорта при оценке производительности команды и отдельных игроков во время тренировок и/или официальных игр [1]. Под отслеживанием позиций игрока во время тренировки понимается вычисление расположения каждого игрока на игровой площадке с достаточной точностью и частотой [2-3], чтобы можно было рассчитать информацию о пути, такую как расстояние, профиль скорости и ускорение.

Подобные данные являются основой для дальнейшего анализа более высокого уровня, такого как стратегия, производительность и физическая подготовка игрока, и помогают тренерам разрабатывать лучшие модели тренировок. Исследования в области захвата движений человека во время спортивных состязаний (тренинги или игры) начались примерно в середине 1920-х годов. Спортивные эксперты всегда заинтересованы в получении данных о позициях игроков во время спортивного матча. Такая информация позволит им анализировать тактическое поведение, действия определенных игроков или всей команды. Ранние исследования в этой области требовали записи спортивного события (матча или тренировки) с использованием видеокамеры, а затем многочасового утомительного ручного ввода, чтобы получить позиции некоторых (не всех) игроков.

Задача обнаружения игрока на спортивном видео в последние годы решается различными способами. Наиболее распространенные подходы основаны на вычитании фона. Например, Zhong et al. [4] представили независимый метод глобальной цветовой фильтрации для выделения областей игрока из фона. В свою очередь Chang et al. [5] сначала оценивают доминирующий цвет площадки, а затем выявляют кандидатов-игроков на заднем плане. Эти подходы эффективны, но на их эффективность могут легко повлиять изменения освещенности, движения камеры и присутствие зрителей [6]. Чтобы сделать обнаружение игроков более надежным, исследователи разработали более сильные функции, такие как гистограмма ориентированных градиентов (HOG) [7], с помощью метода опорных векторов (SVM). Кроме того, исследователи объединили различные функции, такие как ребро LBP и движение, или использовали модель на основе частей [1] для улучшения производительности.

Спортивное сообщество, заинтересованное в получении игровых траекторий во время игры, не ограничивается профильными аналитиками. Тренеры также могут использовать эту информацию, чтобы индивидуализировать планы физической подготовки своих игроков и разработать подходящие наступательные/оборонительные стратегии для достижения наилучших результатов.

Из вышесказанного можно сделать вывод, что отслеживание игроков является основой дальнейшего анализа, который интересен спортивным экспертам. Предлагаемая система слежения основана на цифровой камере, установленной на потолке спортивного зала Ярославского государственного университета им. П.Г. Демидова. Расположение камеры на высоте 6,5 метров с вертикальным ракурсом вниз позволяет получить изображение сразу всего баскетбольного поля. Используемый здесь метод трекинга – это сопоставление шаблонов с некоторыми улучшениями, основанными на предположениях, взятых из правил игры и среды мониторинга [8-10].

Для построения эффективных алгоритмов спортивной видеоаналитики на основе методов глубокого машинного обучения, компьютерного зрения и сверточных нейронных сетей на данном этапе НИОКР решаются следующие основные задачи:

* разработка алгоритмов детектирования объектов на видеоданных с вертикального ракурса камеры;
* разработка алгоритмов трекинга объектов на видеоданных с вертикального ракурса камеры;
* обучение и тестирование алгоритмов трекинга и детектирования на основе сверточной нейронной сети.

Полученные результаты будут использованы на следующих этапах НИОКР и в коммерческой деятельности ООО «Цифровые решения».

**3. ОСНОВНАЯ ЧАСТЬ**

**3.1. Разработка алгоритмов детектирования объектов на видеоданных с вертикального ракурса камеры.**

**3.1.1. Получение изображений с камеры**

При разработке алгоритмов детектирования и отслеживания баскетболистов следующие утверждения принимались в качестве основополагающих: камера зафиксирована и выходит на игровую площадку; игровая площадка ограничена, и ее модель может быть рассчитана; в заданный момент времени два игрока не могут занимать одну и ту же позицию.

В настоящее время использование видеоданных, полученных при помощи вертикального ракурса камеры, распространено в системах видеонаблюдения и мониторинга [11-12]. Эта же настройка камеры нередко используется при съемке спортивных матчей, в частности баскетбольных [13]. Такие камеры устанавливаются в верхних частях спортивного зала или стадиона, что позволяет получить обзор поля целиком. Другим преимуществом использования камер с вертикальным ракурсом является возможность видеть в кадре всех игроков, так как исключаются ситуации, когда несколько игроков, расположенных перед объективом камеры, заслоняют друг друга, тем самым затрудняя процесс распознавания. Недостатком таких камер может выступать трудность выбора точки обзора, при которой в объектив попадают все области интереса. Так, наиболее предпочтительно монтирование камеры на потолок зала, однако даже в этом случае возможности камеры будут ограничены высотой стен. Нередко для решения данной проблемы используются специализированные объективы (например, типа «рыбий глаз») или размещаются несколько камер в разных частях потолка [14].

Для того чтобы отслеживаемые игроки постоянно находились в поле зрения во время игры, на втором этапе научных работ в спортивном зале Ярославского государственного университета им. П.Г. Демидова была установлена стационарная купольная цифровая камера GV-EVD3100. Выбор данной камеры и ее расположения был сделан после тщательного анализа и тестирования многих типов камер с различными датчиками. Высота установки камеры составляет 6,5 метров.

Используемая камера разработана в соответствии со стандартами IK10+ и IP67, за счет чего может использоваться в исследуемых задачах без риска выведения из строя при попадании спортивными снарядами, такими как баскетбольный мяч. В состав данной камеры входят автоматический ИК-фильтр, ИК-светодиоды, светочувствительная 1/2.8” CMOS-матрица с прогрессивной разверткой. Размещение данной камеры на полотолке спортивного зала возможно благодаря регуляции положения объектива. На рисунке 1 представлена схема расположения камеры в спортивном зале ЯрГУ им. П.Г. Демидова. Пример полученного на ее выходе изображения можно увидеть на рисунке 2.

|  |
| --- |
|  |
| Рисунок 1 – Схематическое расположение купольной камеры в спортивном зале ЯрГУ им. П.Г. Демидова |

|  |
| --- |
|  |
| Рисунок 2 – Пример изображения на выходе купольной камеры |

Используемая камера поддерживает различные разрешения для съемки при соотношении 4:3, 16:9, 5:4. В данной исследовательской работе осуществлялась съемка фрагментов баскетбольной тренировки в спортивном зале при использовании максимального разрешения 2048x1536 пикселей. Фокусное расстояние объектива купольной камеры GV-EVD3100 равен 3-9 мм, максимальная апертура равна F/1.7, а обзор может составлять от 36 до 96 градусов.

Созданные видеопоследовательности были использованы для формирования обучающей и тестовой баз данных.

**3.1.2. Обнаружение фона**

Удаление фона – это базовая операция в компьютерном зрении, при которой каждый пиксель в текущей сцене классифицируется как фон или передний план, так что обработка направлена на объекты переднего плана  
 [4-5]. Чтобы реализовать такую сегментацию, должна быть вычислена или получена модель фона. Это можно сделать, сохранив кадр пустого поля. Однако для разных условий освещения может потребоваться обновление фонового изображения, что может быть затруднено во время игры. Решением проблемы является вычисление фоновой модели во время или до начала процесса анализа и, при необходимости, возможность обновления модели во время анализа.

Предлагаемая в данной работе система отслеживания позволяет использовать ранее сохраненное пустое фоновое изображение (рисунок 3). Также она может создать фоновую модель из нескольких последовательных кадров до начала процесса отслеживания. Так, алгоритм обнаружения и вычитания фона состоит из следующих операций для каждого входного кадра:

* сглаживание изображения с помощью фильтра Гаусса;
* в случае использования обновления фона:
  + вычисление разницы между текущим и предыдущим изображениями;
  + применение порога и расширение на разницу между изображениями;
  + копирование области, в которой отсутствуют движущиеся объекты, в модель фона;
* вычисление разницы между текущим кадром и моделью фона;
* задание порога разницы.

|  |
| --- |
|  |
| Рисунок 3 – Изображение фона |

**3.1.3. Сопоставление с шаблоном для детектирования игроков на баскетбольной площадке**

Выбор подходящих признаков очень важен для успешного отслеживания. Признаки, используемые для отслеживания, могут представлять собой либо форму цели, либо информацию о ее цвете [1-3]. Для детектирования в данной системе был использован метод сопоставления с шаблоном [15], т.е. подизображением, содержащим форму, которую требуется найти. Шаблоны формируются с использованием простых геометрических фигур или силуэтов. Сопоставление с шаблоном – это один из методов, используемых в области распознавания образов, который применяется для поиска частей изображения, соответствующих шаблонному изображению. Изображение шаблона сравнивается со всеми частями искомого изображения, и мера сходства вычисляется на каждом этапе сравнения. Результатом процедуры сопоставления является матрица, в которой интенсивность каждого пикселя указывает на степень сходства рассматриваемого участка изображения с шаблоном. Позиция наилучшего совпадения определяется с помощью поиска максимального значения интенсивности.

Преимущество шаблона заключается в том, что он несет как пространственную информацию, так и информацию о внешнем виде изучаемого объекта. Тем не менее, шаблоны кодируют только внешний вид объекта, созданный из одного представления. Таким образом, они подходят лишь для отслеживания объектов, позы которых существенно не меняются в ходе отслеживания [16].

В качестве меры сходства при сопоставлении шаблона и анализируемого изображения был использован нормализованный коэффициент корреляции [17], который вычисляется с использованием следующего уравнения:

где

Используемые обозначения: – результирующее изображение, – шаблон, где – координата каждого пикселя в шаблоне, – исследуемое изображение, где – координаты каждого пикселя исследуемого изображения, – ширина шаблона, - высота шаблона.

При отслеживании людей, в частности спортсменов, могут быть использованы некоторые логические правила, которые исходят от физики или логики, для улучшения отслеживания, чтобы сделать его более эффективным. Одним из правил, которое следует учитывать, является предположение, что первый шаблон выбирается вручную перед началом процесса отслеживания, и что игрок не может внезапно исчезнуть из поля зрения камеры или полностью произвольно изменить свою позицию. Использование таких правил облегчает фокусирование поиска в некоторых областях интереса, что сокращает время обработки.

Чтобы использовать скорость игрока, для предсказания возможной позиции в следующем кадре, следует использовать модель движения игроков. Наиболее популярные динамические модели движения игроков: случайное блуждание (RW) и движение с почти постоянной скоростью (NCV) [2-3, 18]. Модель RW лучше всего описывает движение, если оно имеет более резкие радикальные ускорения в разных направлениях, в то время как модель NCV лучше описывает тип, когда игрок постоянно движется в разных направлениях.

**3.1.4. Обнаружение отдельного игрока**

Чтобы использовать сопоставление с шаблоном для обнаружения и отслеживания баскетболистов на игровой площадке, необходимо учитывать некоторые моменты.

1. Просмотр всех возможных ориентаций спортивного игрока во время игры и настройка шаблона для всех возможных ситуаций потребует большого количества шаблонов, что нецелесообразно.
2. Из-за высокой динамики и изменения характера движений спортивных игроков использование статического шаблона не является оптимальным. Таким образом, шаблон должен обновляться, или должно быть несколько шаблонов, используемых для отслеживания.
3. Размер шаблона для игрока сложно задать фиксированным значением, что связано с настройкой камеры и передвижением баскетболиста по полю в течение матча. Данная особенность заметна на рисунке 4, где показано расположение одного и того же игрока в разные моменты времени на разных позициях, попадающих в обзор камеры. Заметно, что размер тела и ориентация игрока меняются. Рисунок 5 показывает снимок, полученный используемой в исследовании камерой во время одной из тренировок в спортивном зале ЯрГУ им. П.Г. Демидова.

|  |
| --- |
|  |
| Рисунок 4 – Совокупность кадров, на которых игрок занимает различные позиции на игровом поле, полученная при помощи купольной камеры |

Предлагаемое решение проблемы разных размеров и ориентаций игроков состоит в том, чтобы отслеживать верхнюю часть игрока, которая является головой и частью плеч [19]. На рисунках 4 и 5 можно заметить, что головы игроков видны во всех направлениях, и их круглая форма не сильно меняется, как и форма всего тела. Кроме того, при обучении алгоритма Преимущество отслеживания головы заключается в том, что вращение или масштабирование головы не требуется. Таким образом, отслеживание головы и небольшой части плеч баскетболистов проще, чем отслеживание всего тела. В свою очередь, обнаружение головы игрока облегчает расчет положения ног. Хотя верхние камеры уменьшают вероятность того, что голова игрока полностью невидима из-за окклюзии, в некоторых ситуациях, когда игрок падает или другой высокий игрок вступает с ним в сильный контакт, голова исчезает со сцены.

Размер шаблона выбирается в зависимости от разрешения изображения и на практике должен задаваться перед баскетбольным матчем до начала процесса отслеживания.

|  |
| --- |
|  |
| Рисунок 5 – Снимок площадки спортивного зала ЯрГУ им. П.Г. Демидова во время баскетбольной тренировки, полученный купольной камерой |

На рисунке 6 показан шаблон для обнаружения головы баскетболиста, его размеры 25 × 25 пикселей. Рисунок 7 демонстрирует результат процесса сопоставления с шаблоном. Изображения слева – это искомые изображения. Изображения во втором столбце являются используемыми шаблонами. Третий столбец показывает количество лучших совпадений.

|  |
| --- |
|  |
| Рисунок 6 – Пример шаблона головы игрока |

|  |
| --- |
|  |
| Рисунок 7 – Пример результата сопоставления шаблона с игроком в разные моменты времени |

**3.1.5. Обнаружение нескольких игроков**

Предположим, что отслеживаются *N* игроков, и в определенный момент времени позиция каждого игрока известна. Можно разбить изображение на *N* непересекающихся областей, так что в каждой области будет только один игрок. Такой вид деления называется диаграммой Вороного. В данной системе была использована реализация библиотеки Map Manager [20]. На рисунке 8 показан пример диаграммы Вороного для отслеживания нескольких игроков в баскетбольном матче. Это деление используется, чтобы создать маску для каждого из отдельных трекеров и избежать путаницы, когда два или более игрока взаимодействуют очень тесно. Чтобы уменьшить вычислительные затраты на создание таких масок, рассчитывается матрица расстояний всех позиций игрока, и на основе определенного порога вычисляется маска Вороного. Это означает, что игроки, которые не достаточно близки, чтобы повлиять на других, исключены из расчетов маски Вороного.

|  |
| --- |
|  |
| Рисунок 8 – Использование диаграммы Вороного для отслеживания нескольких игроков |

**3.2 Разработка алгоритмов трекинга объектов на видеоданных с вертикального ракурса камеры**

**3.2.1 Система трекинга**

Система компьютерного зрения может рассматриваться как система, состоящая из нескольких этапов: получение изображения с камеры, обработка изображения, анализ сцены и визуализация результата [2-4].

Этап обработки изображения (также называемый предварительной обработкой) - это этап, на котором изображение улучшается, например, путем фильтрации шума или преобразования цвета [21]. Обработка изображений также позволяет найти важные элементы или части сцены, к которым должен применяться анализ. На этапе анализа сцены используются различные методы и алгоритмы, в том числе методы машинного обучения, для распознавания, отслеживания или реконструкции трехмерной модели объектов на сцене. Заключительным этапом может быть визуализация и представление результатов. За счет обратной связи от этапа анализа сцены к этапу обработки изображения передается некоторая информация, которая может помочь оптимизировать этап обработки изображения.

Отслеживание является типичной проблемой компьютерного зрения [22]. На рисунке 9 показаны шаги отслеживания игроков с помощью компьютерного зрения.

В блоке обработки изображений осуществляется обнаружение и вычитание фона, нахождение областей переднего плана изображения путем сегментирования на изображении движущихся и неподвижных частей [23-25].

На этапе анализа сцены происходит отслеживание игроков с использованием специальных алгоритмов. Чтобы получить положение игроков в метрах на игровом поле, необходимо преобразование координат из области изображения в область реального пространства.

|  |
| --- |
|  |
| Рисунок 9 – Этапы системы компьютерного зрения для отслеживания игроков в командных видах спорта. |

Алгоритм отслеживания использует метод на основе сопоставления с шаблоном для отдельного игрока. Результатом трекинга должны быть координаты стопы в метрах, которые также называются преобразованием в координаты реального мира. На первом шаге происходит обнаружение и вычитание фона, за счет чего система получает начальный кадр и инициализирует шаблон. Для каждого входного кадра осуществляется следующая последовательность действий:

* предварительная обработка кадра;
* вычитание кадра из фоновой модели, для нахождения маски переднего плана;
* выполнение сопоставления с шаблоном;
* преобразование выходных координат соответствия шаблона в реальные координаты.

Представленный алгоритм не учитывает отслеживание нескольких игроков, но может выполнять отдельно для каждого игрока.

Для измерения точности отслеживания был создан эталонный набор данных. Полученные траектории от ручного отслеживания используются в качестве истины для дальнейших экспериментов [26]. Размету траектории не требовалось выполнять для всего матча: достаточно нескольких минут игры при условии, что в нее войдут различные ситуации, такие как быстрая атака, спринты и так далее. Данный набор обучающих изображений был взят из видеозаписи баскетбольной тренировки, записанной в спортивном зале ЯрГУ им. П.Г. Демидова. Результирующий объем выборки составил 3100 кадра. Положения головы, тела и ног игроков отмечались вручную на каждом кадре тестовой базы (рисунок 10).

|  |
| --- |
|  |
| Рисунок 10 – Пример разметки положения головы, тела и ног игрока, созданной вручную |

**3.2.2 Исправление искажений купольной камеры**

Видео, используемое для визуализации процедуры отслеживания положения игроков на спортивном поле, должно быть предварительно обработано из-за искажения, вызванного линзами типа «рыбий глаз». Преобразование из искаженного в неискаженное изображение основано на следующих уравнениях:

где – это расстояние до центра выровненного изображения (Defish), – это расстояние до центра исходного искаженного изображения (Fisheye), а – фокусное расстояние. Все перечисленные величины измеряются в пикселях.

Данное преобразование основано на предположении, что линза сферическая, а искажение только радиальное, а не тангенциальное. Поскольку исправленное изображение больше искаженного, требуется интерполяция. Для устранения интерполяции можно использовать обратное отображение, чтобы повторить некоторые пиксели из исходного искаженного изображения в исправленном. В данной реализации использовались таблицы поиска, где отображение для каждой камеры вычисляется только один раз. Чтобы преобразовать пиксель c местоположением искаженного изображения в пиксель с местоположением неискаженного изображения, мы используем следующие уравнения:

где центр искаженного изображения находится в пикселе , а центр неискаженного изображения .

**3.2.3 Трекинг игрока**

Результатом работы алгоритма отслеживания являются координаты расположения игрока на искаженном изображении с купольной камеры, которые в дальнейшем преобразовываются в координаты реального игрового поля в метрах. Эти данные можно использовать для получения полезной информации, такой как пройденное игроком расстояние, скорость и ускорение [27-28]. Для этого требуется заменить фокусное расстояние на высоту камеры в уравнении (4):

где величины , , , измеряются в метрах, ( – расположение проекции головы игрока на поле.

На рисунке 11 показаны расположения камеры и игрока. Из треугольника на рисунке, можно записать геометрическое уравнение:

где– высота камеры в метрах, – высота игрока в метрах.

Для вычисления положения ног, решается уравнение (13) для , тогда:

где – расположение ног игрока на поле в метрах,

При подстановке из уравнения (11):

Аналогичным образом вычисляется координата .

Вместо того, чтобы находить положение ступней за два этапа через положение проекции головы и затем через положение ступней , можно вычислить данное значение путем вычитания высоты каждого игрока из высоты для объекта. Это означает, что уравнение (9) будет записано следующим образом:

где измеряется в метрах.

Если заменить на в (11), мы получим в (17) координату (в метрах), которая является точно такой же, как в (15), но при этом будет произведено меньше вычислений.

Аналогичным образом вычисляется координата .

|  |
| --- |
|  |
| Рисунок 11 – Вычисление реальных координат игрока на игровом поле |

Разработанная видеосистема отслеживания баскетболистов была протестирована в спортивном зале Ярославского университета им. П.Г. Демидова (рисунок 12).

|  |
| --- |
|  |
| Рисунок 12 – Пример интерактивной визуализации траектории игрока на основе обработки видеоданных алгоритмом |

Один кадр обрабатывается примерно за 100 мс. Дополнительное время может быть затрачено на исправление ошибок в отслеживании, которые возникают из-за небольшого скольжения шаблона в сторону от головы игрока или из-за сложного взаимодействия между игроками, которого невозможно избежать. Другим источником ошибок является неточности при обнаружении фона, что может возникнуть из-за изменения условий освещения во время игры.

**3.3 Обучение и тестирование алгоритмов трекинга и детектирования на основе сверточной нейронной сети**

Ошибки в системе отслеживания возникают, когда игроки двигают своими конечностями или движутся вертикально (например, при прыжках). Алгоритм отслеживания пытается отследить часть тела игрока или всего тела. Например, отслеживание соответствия шаблонов, реализованное в этой работе, является отслеживанием на основе головы. Трекер может не соответствовать голове точно. Это происходит из-за столкновений между игроками, которые могут привести к частичной или полной окклюзии отслеживаемой части. Другим источником ошибки является неточная оценка маски фона. Это может привести к тому, что отслеживаемая часть игрока рассматривается как фон и замаскирована из результатов сопоставления [29-30].

Были взяты два набора тестов. Первый набор включал в себя баскетболистов из разных игр, поэтому нет никакой возможности загрязнить обучающий набор данных информацией из тестового набора. Второй набор включал игроков из той же игры, но не из тех же изображений, которые использовались в тренировочном процессе. Помимо 1000 положительных, было также использовано 1000 отрицательных примеров. Эти изображения в основном содержали наружные сцены без людей. Оценка основана на стандартизированном протоколе под названием Вероятность правильной позы (PCP), который измеряет процент правильно локализованных частей тела. Предсказанная область считается правильной, если ее пересечение с фактической площадью как минимум 60%. В противном случае это ложно-положительное обнаружение. Если модель прогнозирует, что большее количество областей перекрывается с одной фактической областью, то только один прогноз считается правильным, а остальные считаются ложноположительными. В данной работе результаты определения правильной позы показаны на рисунке 13. Из приведенного рисунка видно, что точность определения правильной позы увеличивается по мере того, как число в тренировочном наборе приближается к 1100 примерам. Дальнейшее увеличение размера тренировочного набора не влияет на точность обнаружения, которая сохраняет значение почти 0,822 (82,2%). Тестирование с обоими наборами тестов (первое, которое включало игроков из одной и той же игры, и второе, которое содержало только игроков из разных игр), дало почти идентичные результаты. Первый тестовый набор дал точность 82,1%, в то время как второй дал точность 82,2%. Причина в использовании гистограммы ориентированных градиентов. Данное исследование подтверждает, что предложенный алгоритм может быть успешно применен для обнаружения объектов во время спортивных мероприятий.

Предыдущие результаты были получены путем применения созданных моделей к тестовым изображениям. Эти изображения вырезаны из материалов баскетбольного матча, и в большинстве случаев они содержат только одного игрока. Тестирование изученной модели с таким набором данных не может представить ее полный потенциал. Чтобы получить доступ к полномасштабным возможностям, была использована еще одна процедура тестирования. Создан тестовый набор, состоящий из 100 изображений. Из таблицы 1 видно, что из 973 игроков разработанная модель правильно распознала 760 игроков, что составляет 78,1%. Обнаружение было помечено как правильное, если оно перекрывается с правильным расположением игрока с точностью не менее 60%. Из таблицы также заметно, что модель ошибочно признавала некоторые районы баскетболистами.

|  |
| --- |
|  |
| Рисунок 13 – Точность процесса определения положения (в процентах) в зависимости от размера тренировочного набора |

Таблица 1

Точность обнаружения игрока

|  |  |  |  |
| --- | --- | --- | --- |
| **Модель** | **Распознано** | **Ложноположительные** | **Не распознано** |
| Предложенная модель | 760 | 123 | 213 |

Это ложноположительные результаты, в основном это люди из зала и судьи. Если исключить эти два значения, то будет получены результаты, приведенные в таблице 2. Из таблицы 2 видно, что количество ложноположительных результатов было сокращено почти в четыре раза. Остальные ложноположительные обнаружения – это, в основном, множественные обнаружения (алгоритм распознает некоторых игроков более одного раза).

**4. ЗАКЛЮЧЕНИЕ**

К основным результатам, полученным на втором этапе НИОКР могут быть отнесены следующие:

* Для создания системы спортивной аналитики использовалась антивандальная купольная камера GV-EVD3100, разработанная согласно стандартам IK10+ и IP67. Данная камеры была настроена на вертикальный ракурс съемки, позволяющий отображать в кадре расположения верхних частей туловища и головы каждого игрока без перекрытия одних игроков другими. Монтирование камеры производилось на высоте 6,5 метров от пола.
* В результате съемки данной камерой нескольких фрагментов баскетбольных тренировок в спортивном зале ЯрГУ им. П.Г. Демидова была создана базы обучающих данных размером 1100 изображений. Разрешение каждого изображения составляло 2048x1536 пикселей.
* В результате съемки данной камерой нескольких фрагментов баскетбольных тренировок в спортивном зале ЯрГУ им. П.Г. Демидова была создана базы тестовых данных размером более 2000 изображений. Разрешение каждого изображения составляло 2048x1536 пикселей.
* Был разработан алгоритм детектирования баскетболистов на спортивной площадке, основанный на использовании метода сравнения анализируемого кадра видеоданных с вертикального ракурса купольной камеры и шаблона. Размер шаблона, выбранный для данной работы, равен 25х25 пикселей. В качестве меры сходства изображения и шаблона использовался нормализованный коэффициент корреляции.
* Был предложен алгоритм отслеживания баскетболистов основанный на основе использования метода сравнения с шаблоном и преобразования координат изображения с вертикального ракурса камеры в реальные координаты баскетболиста на спортивной площадке.
* Было проведено тестирование разработанных алгоритмов детектирования и трекинга баскетболистов на баскетбольной площадке. Тестовые изображения классифицировались на распознанные, нераспознанные и ложноположительные в зависимости от результата обработки системой. Результаты определения позиций игроков таковы, что из 973 игроков разработанная модель правильно распознала 760 игроков, следовательно, точность составляет 78,1%. Кроме того, было выявлено, что увеличение обучающей базы больше 1100 изображений не приводит к улучшению результата распознавания.

Сфера использования данной группы алгоритмов – видеоаналитика баскетбольных командных матчей с целью улучшения спортивных характеристик игроков, которая может проводиться различными спортивными организациями.

Работы по второму этапу НИОКР были выполнены в полном объеме в соответствии с техническим заданием и календарным планом. Полученные результаты будут использованы на следующих этапах НИОКР и в коммерческой деятельности ООО «Цифровые решения».

# 5. СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Lu W.L., Ting J.A., Murphy K.P., Little J.J.: Identifying Players in Broadcast Sports Videos Using Conditional Random Fields. // CVPR, 2011.

2. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2005. – 621 с.

3. Форсайт Д.А., Понс Д. Компьютерное зрение. Современный подход. – М.: Вильямс, 2004. – 928 с.

4. Chang Zh., Chang S.F.: Real-time view recognition and event detection for sports video // Journal of Visual Communication and Image Representation, 2004. V. 15(3). P. 330–347.

5. Chang M.H., Tien M.C., Wu J.L.: WOW: wild-open warning for broadcast basketball video based on player trajectory // ACM MM, 2009.

6. Manafifard M., Ebadi H., Moghaddam H.A.: A survey on player tracking in soccer videos // CVIU, 2017. P. 19–46.

7. Dalal N., Triggs B.: Histograms of oriented gradients for human detection // CVPR, 2005.

8. Методы компьютерной обработки изображений / Под ред. В.А. Сойфера. – М.: Физматлит, 2001. – 784 с.

9. Прэтт У. Цифровая обработка изображений. – М.: Мир, 1982. – 479 с.

10. Nixon V.M.S., Aguado A.S.: Feature Extraction and Image Processing // Newnes publisher, 2002. P. 164- 168,

11. Kratz L., Nishino K.: Tracking pedestrians using local spatiotemporal motion patterns in extremely crowded scenes // IEEE Trans. Pattern Anal. Mach. Intell., 2012. V. 34, N. 5. P. 987–1002.

12. Angelova A., Krizhevsky A., Vanhoucke V., Ferguson A.O.D.: Real-time pedestrian detection with deep network cascades // BMVC, 2015.

13. Lifang W., Xiuli H., Hao C., Wei S.: Shot segmentation and classification in basketball videos // Intelligent information hiding and multimedia signal processing, 2007. P. 539–542.

14. Breitenstein M., Reichlin F., Leibe B., Koller-Meier E., Van Gool L.: Online Multi-Person Tracking-By-Detection from a Single Uncalibrated Camera // PAMI, 2010. P. 99.

15. Bishop C.M.: Pattern Recognition and Machine Learning. – Springer, 2006.

16. Lin Z., Davis L.S.: Shape-based human detection and segmentation via hierarchical part-template matching // Pattern Analysis and Machine Intelligence, 2010. V. 32(4). P. 604–618.

17. Lewis J. P.: Fast normalized cross-correlation // Proceedings of Vision Interface 95, Canadian Image Processing and Pattern Recognition Society, 1995. P. 120-123.

18. Лукьяница А.А., Шишкин А.Г.: Цифровая обработка видеоизображений. – М.: Ай-Эс-Эс Пресс, 2009. – 518 с.

19. Zhao T., Nevatia R.: Tracking multiple humans in crowded environment // CVPR2004: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, 2004. P. 406–413.

20. Manly K.F., Cudmore Jr., Meer R.H.: Map Manager QTX, cross-platform software for genetic mapping. JM // Mammalian Genome 12, 2001. P. 930-932.

21. Ekin A., Tekalp A.M.: Robust dominant color region detection with applications to sports video analysis // International conference on image processing, 2003.

22. Alahi A., Boursier Y., Jacques L., Vandergheynst P.: Sport players detection and tracking with a mixed network of planar and omnidirectional cameras // ACM/IEEE International conference on distributed smart cameras, 2009. P. 1–8.

23. Parisot P., Vleeschouwer C. D.: Scene-specific classifier for effective and efficient team sport players detection from a single calibrated camera // CVIU, 2017. P. 74–88.

24. Girdhar R., Gkioxari G., Torresani L., Paluri M., Tran D.: Detect-and-track: Efficient pose estimation in videos // IEEE Conf. on Computer Vision and Pattern Recognition, 2018. P. 350–359.

25. Henschel R., Leal-Taixé L., Cremers D., Rosenhahn B.: Fusion of head and full-body detectors for multi-object tracking // IEEE Conf. on Computer Vision and Pattern Recognition Workshops, 2018. P. 1509.

26. Fragkiadaki K., Arbel´aez P., Felsen P., Malik J.: Learning to segment moving objects in videos // Proc. Conf. Computer Vision and Pattern Recognition (CVPR), 2015. P. 4083 - 4090.

27. Ramanathan V., Huang J., Abu-El-Haija S., Gorban A., Murphy K., Fei-Fei L.: Detecting events and key actors in multi-person videos // IEEE Conf. on Computer Vision and Pattern Recognition, 2016. P. 3043–3053.

28. Maresca M. E., Petrosino A.: Clustering local motion estimates for robust and efficient object tracking // Proc. European Conference on Computer Vision, 2014. P. 244–253.

29. Cao Z., Simon T., Wei S.-E., Sheikh Y.: Realtime multi-person 2d pose estimation using part affinity fields // IEEE Conf. on Computer Vision and Pattern Recognition, 2017. P. 1302–1310.

30. Xiao F., Lee Y.J.: Track and segment: An iterative unsupervised approach for video object proposals // Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016. P. 933 – 942.