

УДК 004.855.5

Обзор методов выявления характерных кадров из видеозаписи футбольного матча

Вершинин Евгений Владимирович

vershinin@bmstu.ru

Сафин Павел Римович

spr20ki167@student.bmstu.ru

КФ МГТУ им. Н.Э. Баумана, Калуга, Россия

Исследование посвящено методу автоматизированного отбора ключевых кадров из видеозаписей футбольных матчей с использованием искусственного интеллекта. Рассмотрены современные подходы и архитектуры нейронных сетей, включая CNN, LSTM и трансформеры, которые позволяют эффективно выявлять моменты, представляющие интерес для зрителей, такие как голы, атаки и фолы. Освещены этапы предварительной обработки и постобработки данных, повышающие точность и скорость анализа. Представленный метод позволяет автоматически генерировать нарезки лучших моментов, сокращая время на анализ и улучшая восприятие зрителями спортивных событий.

Ключевые слова: искусственный интеллект, спортивное видео, нейронные сети, компьютерное зрение, CNN, трансформеры

С развитием технологий и популяризацией спортивных мероприятий наблюдается значительный рост интереса к анализу и обработке видеозаписей спортивных событий, таких как футбольные матчи. Традиционные методы ручного отбора ключевых моментов, таких как голы, опасные моменты и зрелищные атаки, требуют значительных временных и человеческих ресурсов. Это делает задачу выделения удачных кадров для последующего анализа или показа особенно актуальной. Искусственный интеллект (ИИ) предоставляет возможности для автоматизации и оптимизации процесса выбора таких моментов, делая ее более быстрым и точным.

Цель данной статьи заключается в рассмотрении метода автоматизированного выявления удачных кадров во время видеозаписи футбольного матча с использованием ИИ. Будут рассмотрены основные подходы, архитектуры нейронных сетей, методы предварительной обработки данных и ключевые метрики, позволяющие улучшить качество отбора.

Главная задача метода заключается в выявлении кадров, которые представляют интерес с точки зрения зрелищности или ключевых событий: голы, атаки, фолы, спасения вратаря и другие важные моменты. Для достижения этой цели необходимо разработать алгоритм, который может:

- распознавать основные события матча, включая голы, опасные атаки, пенальти и т. д.;
- выделять значимые кадры для каждого события, избегая пропуска важных моментов;
- поддерживать высокую точность и производительность для анализа больших объемов видеозаписей в реальном времени или близки к нему.

Первым шагом является предварительная обработка данных. Необходимо сегментировать видеозапись на короткие отрезки и анализировать каждый кадр или серию кадров. Этот этап включает следующие подзадачи:

- распознавание объектов — выделение футболистов, мяча, ворот и других элементов поля;
- отслеживание объектов — идентификация траекторий движения ключевых объектов (например, игроков и мяча) на протяжении времени;
- распознавание поз — анализ поз игроков, который помогает понять их действия, такие как удары по мячу или защитные маневры.

Для этих задач часто применяют сверточные нейронные сети (CNN) и сети глубокого обучения для задач компьютерного зрения, такие как ResNet, MobileNet или YOLO, которые обеспечивают высокую производительность и точность при идентификации объектов на изображениях [1].

Поскольку футбольный матч представляет собой последовательность действий, важно учитывать временные зависимости между кадрами. Здесь подходят рекуррентные нейронные сети (например, LSTM и GRU), которые способны запоминать и обрабатывать последовательности данных, анализируя контекстное значение каждого кадра [2]. Эти сети можно использовать для анализа последовательностей изображений, выделяя моменты с высокими вероятностями значимых событий. Например, если система фиксирует внезапное движение группы игроков в сторону ворот, а также траекторию мяча, она может оценить вероятность потенциального гола.

Трансформеры, такие как Vision Transformer (ViT), все чаще применяются в анализе спортивных видео [3]. Их способность учитывать контекст на уровне целого события делает их особенно полезными для автоматического выявления ключевых моментов. Трансформеры позволяют быстрее и точнее обрабатывать большие видеоданные, выявляя важные моменты с учетом всех игроков на поле.

Для обеспечения точности выявленных событий применяется пост-обработка, включающая фильтрацию ложных срабатываний и улучшение качества кадров. Пост-обработка позволяет отсекать нерелевантные моменты и оптимизировать выборку, сохраняя только те кадры, которые представляют реальный интерес для зрителей [4].

Для оценки работы алгоритма можно использовать следующие метрики:

- точность (Precision) — отношение правильно отобранных кадров к общему числу отобранных кадров;
- полнота (Recall) — отношение правильно отобранных кадров к общему числу релевантных кадров;
- F1-мера — гармоническое среднее точности и полноты;
- плавность переходов — проверка, чтобы переход между кадрами был естественным и не нарушал целостность восприятия событий.

Метод автоматического выделения удачных кадров позволяет значительно сократить время на анализ видеозаписей и выделение ключевых событий. С помощью ИИ можно:

- автоматически генерировать нарезки лучших моментов для последующих аналитических разборов;
- ускорить процесс создания хайлайтов и обзорных роликов;
- повысить зрелищность и удобство просмотра матчей для болельщиков.

Тестирование предложенного метода на реальных видеозаписях футбольных матчей показало, что использование сверточных нейронных сетей и трансформеров позволяет достичь высокой точности и полноты выявления ключевых моментов [5]. Кроме того, производительность алгоритма позволяет работать с видео в реальном времени, что открывает перспективы для использования метода в прямых трансляциях.

Таким образом, использование искусственного интеллекта для автоматизированного отбора удачных кадров в видеозаписях футбольных матчей позволяет упростить и ускорить процесс анализа спортивных событий. Применение методов компьютерного зрения, нейронных сетей и трансформеров помогает эффективно выделять значимые моменты, что делает просмотр спортивных событий более увлекательным для зрителей.

Литература

- [1] Скрипальщиков И.И. Анализ алгоритмов компьютерного зрения обнаружения движения. *Студ. науч.-техн. конф., посв. 100-летию отечественной гражданской авиации: сб. докл.* Москва, ИД Академии Н.Е. Жуковского, 2023, с. 82–86.
- [2] Анисимовский В.В., Пенкин М.А., Дорохов Е.А., Груздев А.М., Завалишин С.С., Курилин И.В. Слепое устранение размытия изображения с использованием многозадачной сверточной нейронной сети и пространственно-вариативной рекуррентной нейронной сети. *ГрафиКон. Междунар. конф. по компьютерной графике и зрению: сб. тр.* Москва, ФИЦ Институт прикладной математики им. М.В. Келдыша РАН, 2023, № 33, с. 455–462.
- [3] Князь В.В., Мизгинов В.А. DiNoNAT: исследование самостоятельного обучения нейронных сетей с использованием трансформеров и механизмов локального внимания. *Моделирование авиационных систем. V Всерос. науч.-техн. конф.: сб. тез. докл.* Москва, Государственный научно-исследовательский институт авиационных систем, 2023, с. 322–323.
- [4] Садовая А.С. Графическая пост-обработка изображений. *Вестник молодых ученых Санкт-Петербургского государственного университета технологии и дизайна*, 2020, № 4, с. 51–55.
- [5] Дудин Э.С., Понкратов Е.С. Оценка производительности предобученных сверточных нейронных сетей для задач классификации изображений. *Актуальные проблемы авиации и космонавтики. IX Междунар. науч.-практ. конф., посв. Дню космонавтики: сб. материалов.* Красноярск, Сибирский государственный университет науки и технологий им. академика М.Ф. Решетнева, 2023, с. 166–168.

Review of methods for identifying characteristic framework from a video recording of a football match

Vershinin E.V.

vershinin@bmstu.ru

Safin P.R.

spr20ki167@student.bmstu.ru

BMSTU (KB), Kaluga, Russia

This study is devoted to the method of automated keyframe extraction from football video recordings using artificial intelligence. The paper considers modern approaches and neural network architectures, including CNN, LSTM and transformers, which allow to effectively identify moments of interest to viewers, such as goals, attacks and fouls. It also covers the stages of data pre-processing and post-processing, which increase the accuracy and speed of analysis. The presented method allows to automatically generate highlights, reducing the time for analysis and improving the perception of sports events by viewers.

Keywords: artificial intelligence, sports video, neural networks, computer vision, CNN, transformers