

УДК 004.855.5

## Автоматическое тегирование событий футбольного матча на основе видеоданных

Вершинин Евгений Владимирович

vershinin@bmstu.ru

Сафин Павел Римович

spr20ki167@student.bmstu.ru

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

*Рассмотрена проблема автоматического тегирования событий футбольного матча с использованием нейросетевых моделей. Описаны методы предварительной обработки данных, включая разбиение видеопотока на кадры, нормализацию изображений и аугментацию данных для повышения устойчивости модели. Рассмотрены подходы к распознаванию ключевых моментов, такие как выделение пространственных признаков с помощью сверточных нейросетей и анализ временных зависимостей через рекуррентные сети. В качестве основной архитектуры предложена комбинация ResNet и LSTM, которая позволяет эффективно анализировать футбольные события, включая голы, карточки, офсайды и опасные моменты.*

**Ключевые слова:** автоматическое тегирование, анализ видео, компьютерное зрение, машинное обучение, нейросетевые модели, ResNet, LSTM, спортивная аналитика, футбольные события, распознавание действий

Автоматический анализ спортивных событий является актуальной задачей в области компьютерного зрения и машинного обучения. Одним из наиболее сложных видов контента для анализа является футбольный матч, который характеризуется динамическими изменениями сцены, большим количеством игроков и сложными взаимодействиями [1]. Основная цель данной работы — разработка модели, способной анализировать видеопоток футбольного матча и автоматически присваивать тегированные метки ключевым событиям.

Для обучения модели требуется обширный датасет, содержащий видеозаписи матчей и их разметку по ключевым событиям. Источники данных могут включать:

- открытые датасеты: например, SoccerNet, FIFA Dataset и другие наборы данных, содержащие размеченные события матчей;
- полные трансляции матчей: видеопоток, полученный с различных спортивных каналов или официальных трансляций;
- аннотированные нарезки событий: фрагменты видео, в которых вручную выделены важные моменты (например, голы, офсайды, нарушения).

Разметка данных является важным этапом подготовки датасета. Она может выполняться вручную, когда эксперты просматривают видео и указывают временные метки ключевых событий, или с использованием полуавтоматических инструментов, которые анализируют аудио- и видеопоток для предварительной разметки. В процессе аннотации могут использоваться дополнительные источники данных:

- аудиодорожки матчей: анализ комментариев, шума трибун и голосов судей может помочь точнее определять события (например, повышение громкости комментатора при голе);

- трекары игроков и мяча: использование данных о координатах игроков и мяча может улучшить точность идентификации событий, особенно при определении офсайдов и штрафных ударов [2].

Для улучшения качества обучения модели проводится аугментация данных, включающая изменение яркости, контраста, разрешения, а также добавление искусственных шумов. Это помогает сделать модель устойчивой к различным условиям съемки и освещения [3].

Автоматическое тегирование событий включает несколько этапов, каждый из которых играет важную роль в процессе распознавания ключевых моментов матча [4]:

- извлечение ключевых кадров: использование OpenCV для разбиения видео на последовательности изображений с заданной частотой кадров. Это позволяет выделить наиболее информативные моменты, уменьшить объем обрабатываемых данных и сократить избыточность кадров;

- предобработка изображений: включает изменение размера, нормализацию, а также аугментацию данных (вращение, изменение яркости и контрастности) для улучшения устойчивости модели к вариациям в видеопотоке;

- глубокая сверточная сеть (ResNet): используется для извлечения пространственных признаков из каждого кадра. ResNet обучается на размеченных данных, выявляя характерные паттерны, например, движение мяча, положение игроков, изменение освещенности;

- рекуррентная сеть (LSTM): после извлечения признаков с помощью ResNet они передаются в LSTM, которая анализирует временные зависимости между последовательными кадрами. Это позволяет учитывать контекст матча, что особенно важно при распознавании событий, требующих анализа последовательности действий, например, гола или нарушения;

- классификация событий: финальный этап, на котором полносвязный слой модели присваивает кадрам вероятностные метки соответствующих событий (например, «гол», «угловой», «офсайд»). На данном этапе применяется функция потерь, например, CrossEntropyLoss, для оптимизации предсказаний модели и повышения точности классификации.

Выбранная архитектура основана на комбинации сверточной нейросети ResNet и рекуррентной сети LSTM. Сверточная нейросеть (CNN) отвечает за выделение пространственных признаков на каждом кадре видео, а рекуррентная сеть (LSTM) анализирует временную зависимость между последовательными кадрами. Такая комбинация позволяет учитывать как статические, так и динамические аспекты футбольного матча, что особенно важно при детекции событий, развивающихся во времени.

**Реализация модели.** Предобработка видео:

- разбиение видео на кадры с частотой 1 FPS для уменьшения объема данных и сохранения ключевых изменений сцены;

- изменение разрешения изображений и нормализация значений пикселей для обеспечения единообразия входных данных;

- аугментация данных (изменение яркости, поворот, масштабирование) для повышения устойчивости модели к различным условиям освещения и углам съемки.

Обучение нейросети:

- использование предобученной ResNet-18 с удалением полносвязного слоя и заменой его на новую архитектуру, оптимизированную для задач классификации футбольных событий;

- извлечение признаков с помощью ResNet и передача их в LSTM, которая анализирует последовательность кадров;

- добавление полносвязного слоя с функцией активации softmax для классификации на заданные категории событий (например, «гол», «фол», «офсайд»).

Оптимизация модели:

- функция потерь: CrossEntropyLoss, подходящая для многоклассовой классификации;

- оптимизатор: Adam с динамическим изменением скорости обучения для повышения эффективности обучения;

- использование dropout и batch normalization для предотвращения переобучения;

- балансировка классов в обучающем наборе данных для повышения точности предсказаний на редкие события.

Таким образом, автоматическое тегирование футбольных событий на основе нейросетей — перспективное направление, которое может применяться для создания нарезок матчей, улучшения спортивной аналитики и автоматизации комментариев. В будущем возможно улучшение точности модели за счет мультимодального анализа (видео + аудио) и применения трансформеров для обработки видеопотока.

## Литература

- [1] Амонтурдиев Б.К. Анализ разбора стиля игры команды для контратаки в футболе. *Столица науки*, 2020, № 4 (21), с. 358–362.
- [2] Колчина Е.Ю. Фитнес-трекер как стимулятор физической активности и его влияние на здоровье и нагрузку человека. *Вестник Луганского государственного университета имени Владимира Даля*, 2022, № 8 (62), с. 51–54.
- [3] Ложкин И.А., Дунаев М.Е., Зайцев К.С., Гармаш А.А. Аугментация наборов изображений для обучения нейронных сетей при решении задач семантической сегментации. *International Journal of Open Information Technologies*, 2023, т. 11, № 1, с. 109–117.
- [4] Рябко В.В. Видеоаналитика с использованием сверточной нейронной сети. *Международ. науч.-техн. конф. молодых ученых БГТУ им. В.Г. Шухова, посв. 300-летию Российской академии наук. Нац. конф. с междунар. уч.: сб. докл.* Белгород, БГТУ им. В.Г. Шухова, 2022, с. 357–360.

## Automatic tagging of football match events based on video data

Vershinin Evgeny Vladimirovich

vershinin@bmstu.ru

Safin Pavel Rimovich

spr20ki167@student.bmstu.ru

*Kaluga Branch of BMSTU, Kaluga, Russia*

This paper addresses the problem of automatic tagging of football match events using neural network models. Data preprocessing methods are described, including splitting the video stream into frames, image normalization, and data augmentation to improve the robustness of the model. Approaches to recognizing key moments are considered, such as extracting spatial features using convolutional neural networks and analyzing temporal dependencies through recurrent networks. A combination of ResNet and LSTM is proposed as the main architecture, which allows for efficient analysis of football events, including goals, cards, offsides, and dangerous moments.

**Keywords:** minimal tagging, video analysis, computer vision, machine learning, neural network models, ResNet, LSTM, sports analytics, football events, action recognition