

Sistema de Detección y Seguimiento Automático de Jugadores y Pelota en Partidos de Tenis en Tiempo Real

Marino Fernández Pérez

marinofepe@correo.ugr.es

Francesc Oliver Catany

francescoliver@correo.ugr.es

Pau Bover Femenias

pauboover@correo.ugr.es

Gabriele Ruggeri

gabricross37@correo.ugr.es

Universidad de Granada (UGR) - Visión por Computador

Abstract

En este trabajo se presenta un sistema para el análisis automático de un partido de tenis a partir de vídeo broadcast. El sistema es capaz de identificar la pista de juego, sus líneas y keypoints relevantes, así como detectar y seguir a los jugadores y la pelota a lo largo de la secuencia de vídeo. A partir de esta información, se estiman métricas cinemáticas como la velocidad y la distancia recorrida por los jugadores durante el partido.

La solución propuesta integra múltiples modelos de aprendizaje profundo junto con técnicas clásicas de visión por computador, organizados en un pipeline modular y extensible. El sistema ha sido evaluado en vídeos reales de diferentes competiciones y tipos de pista, mostrando un funcionamiento robusto en escenarios variados. Aunque el procesamiento se realiza actualmente en un régimen de casi tiempo real, se discute su viabilidad para aplicaciones en tiempo real mediante optimizaciones adicionales. Los resultados obtenidos demuestran el potencial del enfoque propuesto como herramienta de análisis automático en el ámbito del tenis profesional.

La resolución de este problema resulta relevante debido al creciente interés en el uso de herramientas basadas en inteligencia artificial para el análisis avanzado del rendimiento deportivo. La información recopilada por sistemas automáticos de detección y seguimiento permite desarrollar métricas complejas y estimaciones sobre el comportamiento de los jugadores durante el partido. Este tipo de análisis puede proporcionar una ventaja competitiva en el estudio de rivales, facilitando la identificación de patrones de juego y tendencias estratégicas. Además, este tipo de tecnología presenta un alto potencial de aplicación en retransmisiones televisivas, donde puede emplearse para ofrecer información visual y estadística de interés que enriquezca la experiencia de los espectadores.

El objetivo de este trabajo es desarrollar un sistema que permita detectar de manera automática información relevante a partir de partidos de tenis, de forma que cualquier usuario pueda utilizarlo para analizar vídeos y construir métricas y análisis avanzados basados en dicha información.

1.1. Motivación personal

Elegimos este tema como proyecto porque nos pareció un campo especialmente amplio y estimulante, en el que era posible profundizar en muchos de los conceptos que más nos habían interesado a lo largo de la asignatura. El análisis de vídeo deportivo combina técnicas de aprendizaje profundo, procesamiento clásico de imágenes y geometría, lo que nos permitió aplicar de forma práctica conocimientos diversos que iremos detallando más adelante. Además, la complejidad real del problema y la necesidad de tomar decisiones de diseño fundamentadas lo convirtieron en un contexto idóneo para consolidar y ampliar lo aprendido durante el curso.

1. Introducción

En el ámbito del tenis profesional, el análisis detallado del rendimiento tanto de jugadores propios como de rivales supone una tarea compleja y costosa para entrenadores y analistas. Gran parte de este análisis se realiza de forma manual o semiautomática, lo que dificulta la obtención de información objetiva y detallada a partir de grandes volúmenes de vídeo.

2. Contexto

El análisis automático de partidos de tenis a partir de secuencias de vídeo constituye un problema relevante dentro del ámbito de la visión por computador, al requerir la detección y el seguimiento de elementos dinámicos de la escena, como jugadores y pelota, así como la localización precisa de estructuras geométricas estáticas, en particular la pista. La combinación adecuada de técnicas de aprendizaje profundo con métodos clásicos de procesamiento de imagen resulta fundamental para abordar estos retos de forma robusta y coherente.

En los últimos años, los modelos basados en aprendizaje profundo han mostrado un alto rendimiento en tareas de detección y seguimiento de objetos en tiempo real. Arquitecturas como YOLO se utilizan habitualmente para la detección de jugadores, mientras que redes especializadas como TrackNet han sido diseñadas para el seguimiento de objetos pequeños y de movimiento rápido, como la pelota. Asimismo, modelos de aprendizaje profundo se emplean para la detección de puntos clave relevantes de la pista, proporcionando estimaciones iniciales que pueden ser posteriormente refinadas. El uso de modelos ligeros permite además realizar inferencia eficiente, facilitando su integración en sistemas de análisis de vídeo deportivo.

Por otro lado, la detección de las líneas de la pista presenta características distintas, ya que se trata de estructuras geométricas bien definidas y con una disposición regular. En este contexto, los métodos clásicos de procesamiento de imagen, como la Transformada de Hough, continúan siendo una alternativa eficaz, especialmente cuando se combinan con técnicas de preprocessado que reducen el ruido y mejoran la calidad geométrica de las detecciones.

Finalmente, la relación entre la imagen capturada por la cámara y el plano real de la pista puede modelarse mediante técnicas de geometría proyectiva. El uso de homografías permite establecer una correspondencia entre ambos planos, posibilitando la superposición de modelos de referencia y la interpretación espacial de las posiciones detectadas. Para garantizar la fiabilidad de esta transformación, resulta esencial disponer de puntos clave precisos y geométricamente consistentes.

3. Trabajos Previos

El análisis automático de partidos de tenis ha sido objeto de un interés creciente en los últimos años, impulsado tanto por el desarrollo de técnicas de visión por computador como por la disponibilidad de modelos de aprendizaje profundo capaces de operar en tiempo real. Los trabajos

previos en este ámbito pueden agruparse, de forma general, en tres grandes líneas: la detección y seguimiento de jugadores, el seguimiento de la pelota y la detección de la pista y sus líneas.

En el ámbito del seguimiento de la pelota, uno de los trabajos más relevantes es *TrackNet* [?], una red neuronal profunda diseñada específicamente para el seguimiento de objetos pequeños y de alta velocidad en aplicaciones deportivas. TrackNet aborda el problema como una tarea de segmentación, generando mapas de probabilidad que indican la posición de la pelota en cada frame. Este enfoque ha demostrado ser especialmente eficaz en deportes como el tenis, donde la pelota ocupa un número muy reducido de píxeles y presenta movimientos rápidos y abruptos. Debido a su robustez y precisión, TrackNet se ha convertido en una referencia ampliamente utilizada en sistemas de análisis de tenis basados en vídeo.

En lo que respecta a la detección de jugadores, numerosos trabajos y aplicaciones prácticas emplean modelos de detección de propósito general basados en aprendizaje profundo, como YOLO o variantes de Faster R-CNN. Estos modelos permiten localizar jugadores de forma eficiente incluso en escenarios complejos, con cambios de iluminación, occlusiones parciales y variaciones de escala. Su uso se ha consolidado como una solución estándar dentro de pipelines de análisis deportivo en tiempo real.

La detección de la pista y de sus líneas presenta características particulares que han dado lugar a distintos enfoques en la literatura. Existen propuestas que aplican técnicas de aprendizaje automático para esta tarea, como el trabajo presentado por ML6 [?], donde se emplean modelos de *machine learning* para mejorar la detección inicial de las líneas de la pista, manteniendo métodos clásicos en fases posteriores del procesamiento. Este tipo de enfoques muestra que las redes neuronales aportan una mayor robustez frente a variaciones de iluminación, sombras u occlusiones parciales, mientras que las técnicas tradicionales permiten preservar la coherencia geométrica del resultado.

En esta línea, destacan implementaciones prácticas y de código abierto como las propuestas por Tarek [?] y Yastrebkov [?]. Estos trabajos adoptan enfoques híbridos que combinan detección basada en aprendizaje profundo con técnicas clásicas de procesamiento de imagen y razonamiento geométrico como la Transformada de Hough. Aunque estos métodos pueden presentar limitaciones en condiciones extremas, ofrecen soluciones eficientes, interpretables y ampliamente utilizadas como referencia en el análisis automático de tenis.

En conjunto, los trabajos previos muestran que los enfoques más eficaces para el análisis automático de partidos

de tenis no se basan en una única técnica, sino en la combinación de modelos de aprendizaje profundo para la detección y el seguimiento de los elementos dinámicos con métodos clásicos y geométricos para el fortalecimiento de la detección de las estructuras estáticas de la pista. Esta integración permite aprovechar las fortalezas de ambos paradigmas y constituye la base de los sistemas más robustos y coherentes propuestos en la literatura.

4. Methods

Detailed description of the methods used and/or proposed, and clear justification of why these methods are used and not others.

5. Experiments

The data used, the experimental validation protocol, the metrics used, the experiments carried out, the results obtained, and their discussion are presented here.

5.1. Dataset

6. Conclusions

Section that presents, briefly and as a summary, the main conclusions of the work carried out. It also usually includes future possible works. That is, what are the most promising lines to continue with this work, as well as possible proposals for improvement. IMPORTANT: these are the scientific conclusions reached in the project; not your personal conclusions about the work you have done!

References