



**Universidad
Europea**

UNIVERSIDAD EUROPEA DE MADRID

ESCUELA DE ARQUITECTURA, INGENIERÍA Y DISEÑO

MASTER UNIVERSITARIO EN ANALISIS DE DATOS MASIVOS (BIG DATA)

TRABAJO FIN DE MÁSTER

**Desarrollo de modelo IA para optimización de
decisiones y rendimiento en partidos de fútbol**

ANGEL MOLPECERES MINGO

Dirigido por

CARLOS JAIME PAGES OGILVIE

CURSO 2024-2025

Angel Molpeceres Mingo

TÍTULO: DESARROLLO DE MODELO IA PARA OPTIMIZACIÓN DE DECISIONES Y RENDIMIENTO EN PARTIDOS DE FÚTBOL

AUTOR: ANGEL MOLPECERES MINGO

TITULACIÓN: MÁSTER EN ANÁLISIS DE DATOS MASIVOS (BIG DATA)

DIRECTOR/ES DEL PROYECTO: CARLOS JAIME PAGES OGILVIE

FECHA: SEPTIEMBRE DE 2025

RESUMEN

Este TFM desarrolla un sistema completo de visión por computador para reconstruir el estado de juego en fútbol a partir de vídeo de transmisión, con el objetivo de democratizar el análisis táctico y generar análisis operativos casi en tiempo real. Aborda la falta de un pipeline unificado que integre detección, seguimiento, identificación de dorsales, calibración de cámara y proyección métrica fiable al campo, reduciendo la dependencia de anotación manual y la degradación en material no profesional. La solución implementa, entre otras técnicas, detección y seguimiento basados en YOLO y un sistema de OCR para leer dorsales, junto con una calibración geométrica que transforma posiciones en coordenadas del terreno de juego. A partir de las trayectorias, se construyen mapas que reparten el espacio del campo entre los jugadores según cercanía e influencia dinámica, se detectan momentos donde un futbolista ocupa o libera espacio para un compañero, se estima el peligro potencial de la jugada y se señalan receptores prometedores sin oposición inmediata. Todo ello en una interfaz 2D funcional aplicada a secuencias SoccerNet. Los resultados evidencian viabilidad con hardware común y software open-source, traduciendo datos visuales en métricas tácticas operativas útiles para la toma de decisiones.

Palabras clave: Visión por computador, YOLO, Fútbol, OCR, Análisis táctico.

ABSTRACT

This Master's Thesis develops a complete computer vision system to reconstruct the game state in football from broadcast video, aiming to democratize tactical analysis and deliver operational insights almost in real time. It tackles the lack of a unified workflow that covers detection, tracking, jersey-number identification, camera calibration, and reliable projection onto the pitch, reducing reliance on manual annotation and performance loss on non-professional footage. The solution implements, among other techniques, YOLO-based detection and tracking and an OCR system to read jersey numbers, together with geometric calibration that converts positions into pitch coordinates. From the player trajectories, the system builds maps that distribute pitch space among players based on proximity and dynamic influence, detects moments when a player occupies or frees space for a teammate, estimates the potential danger of an action, and highlights promising receivers without immediate pressure. All of this is integrated into a functional 2D interface applied to SoccerNet sequences. The results show feasibility on common hardware with open-source software, turning visual signals into operational tactical metrics that support decision-making.

Keywords: Computer vision, YOLO, football, OCR, tactical analysis

AGRADECIMIENTOS

Gracias a mis padres por la educación que me han dado y el tiempo y amor que me han dedicado. A mi hermano por estar conmigo desde el día uno, te admiro mucho. A mis amigos más cercanos, por permitirme pasar tan buenos ratos juntos. A ti Bea, por animarme siempre que lo necesito, por esa ilusión que le pones a las cosas y por no rendirte nunca. Va por vosotros.

TABLA RESUMEN

	DATOS
Nombre y apellidos:	Angel Molpeceres Mingo
Título del proyecto:	DESARROLLO DE MODELO IA PARA OPTIMIZACIÓN DE DECISIONES Y RENDIMIENTO EN PARTIDOS DE FÚTBOL
Directores del proyecto:	Carlos Jaime Pages Ogilvie
El proyecto se ha realizado en colaboración de una empresa o a petición de una empresa:	NO
El proyecto ha implementado un producto: (esta entrada se puede marcar junto a la siguiente)	SI
El proyecto ha consistido en el desarrollo de una investigación o innovación: (esta entrada se puede marcar junto a la anterior)	NO
Objetivo general del proyecto:	DESARROLLO DE MODELO IA PARA OPTIMIZACIÓN DE DECISIONES Y RENDIMIENTO EN PARTIDOS DE FÚTBOL

Tabla 1. Tabla de resumen del proyecto

Índice

RESUMEN	3
ABSTRACT	4
TABLA RESUMEN	6
Capítulo 1. RESUMEN DEL PROYECTO	12
1.1 Contexto y justificación	12
1.2 Planteamiento del problema	12
1.3 Objetivos del proyecto	12
1.4 Resultados obtenidos	12
1.5 Estructura de la memoria	13
Capítulo 2. ANTECEDENTES / ESTADO DEL ARTE	14
2.1 Estado del arte	14
2.2 Contexto y justificación	20
2.3 Planteamiento del problema	21
Capítulo 3. OBJETIVOS	24
3.1 Objetivos generales	24
3.2 Objetivos específicos	24
3.3 Beneficios del proyecto	25
Capítulo 4. DESARROLLO DEL PROYECTO	28
4.1 Planificación del proyecto	28
4.2 Descripción de la solución, metodologías y herramientas empleadas	28
4.3 Recursos requeridos	46
4.4 Presupuesto	46
4.5 Viabilidad	46
4.6 Resultados del proyecto	47
Capítulo 5. DISCUSIÓN	66
Capítulo 6. CONCLUSIONES	68
6.1 Conclusiones del trabajo	68

6.2	Conclusiones personales.....	68
Capítulo 7.	FUTURAS LÍNEAS DE TRABAJO	70
Capítulo 8.	REFERENCIAS.....	71
Capítulo 9.	ANEXO	76

Índice de Figuras

Figura 1. Diagrama de Gantt de la planificación del proyecto	28
Figura 2. Generación de insights tácticos mediante diagramas de Voronoi para el análisis del control espacial del terreno de juego [49]	41
Figura 3. Mapa de control del campo [50]	42
Figura 4. Secuencia de generación de espacio. Elaboración propia	44
Figura 5. Aplicación del pipeline a secuencias de partido: imagen de partido	48
Figura 6. Detección de jugadores en el campo con asignación de bounding boxes.....	49
Figura 7. Reconocimiento de números de dorsal en jugadores detectados.....	49
Figura 8. Recorte y procesamiento de la región del dorsal.....	50
Figura 9. Mejora de nitidez mediante superresolución.	50
Figura 10. Transformación cromática y reducción de complejidad visual.....	51
Figura 11. Reducción de la paleta cromática.	51
Figura 12. Simplificación cromática mediante paleta limitada a ocho colores.....	51
Figura 13. Aislamiento secuencial de los colores dominantes.....	52
Figura 14. Aislamiento de dígitos en dorsales mediante selección de componentes visuales relevantes.....	53
Figura 15. Reconocimiento óptico de caracteres (OCR) para la identificación final del dorsal..	53
Figura 16. Resultados completos del pipeline de reconocimiento de jugadores y dorsales.	54
Figura 17. Excepción en el reconocimiento del dorsal.....	55
Figura 18. Identificación de puntos clave en el terreno de juego para la estimación geométrica de la posición real de los jugadores [51].....	55
Figura 19. Detección automática de líneas del campo para validar la localización espacial de los jugadores.....	56
Figura 20. Proyección de posiciones reales de los jugadores sobre el plano en dos dimensiones del campo mediante homografía.....	57
Figura 21. Evolución posicional del jugador ID 22 en secuencia táctica: apertura, rotación y desmarque.	58
Figura 22. Distribución del control territorial durante una jugada ofensiva: concentración defensiva y apertura estratégica.....	59

Figura 23. Evolución de las zonas de control durante una jugada: transición de distribución amplia a concentración táctica.	60
Figura 24. Disputa de posesión y aprovechamiento del espacio tras recuperación por parte del equipo azul.	61
Figura 25. Visualización de eventos SOG activos y pasivos durante el inicio de un contraataque del equipo negro.	62
Figura 26. Generación de espacio mediante atracción defensiva y desmarque en jugada ofensiva del equipo rojo.	63
Figura 27. Visualización de carreras de valor en jugada ofensiva mediante modelo EPV: generación de oportunidades y finalización.	64
Figura 28. Identificación de opciones de pase mediante modelo de ventaja sin balón en jugada ofensiva del equipo blanco.	65

Índice de Tablas

Tabla 1. Tabla de resumen del proyecto	6
Tabla 2. Formato de anotaciones en archivos de ground truth según el estándar MOT Challenge.	29
Tabla 3. Estructura de las detecciones de jugadores por frame	36
Tabla 4. Coordenadas reales de los jugadores en el plano del campo tras la transformación por homografía	40
Tabla 5. Presupuesto del proyecto.....	46

Capítulo 1. RESUMEN DEL PROYECTO

1.1 Contexto y justificación

El proyecto se sitúa en la transición del análisis futbolístico desde la observación subjetiva hacia un enfoque objetivo basado en IA y visión por computador aplicada a vídeo de transmisión, con el propósito de democratizar el acceso a análisis táctico avanzado y generar insights accionables cercanos al tiempo real. Se propone un sistema integral que detecta, identifica y mapea a jugadores al plano del campo para transformar datos visuales en conocimiento táctico útil para la toma de decisiones.

1.2 Planteamiento del problema

La problemática central es la ausencia de un pipeline unificado y accesible que integre detección, tracking, reconocimiento de dorsales, calibración y análisis táctico; además, existe degradación al transferir modelos a vídeo amateur y alta dependencia de anotación manual costosa. El trabajo investiga y prototipa una solución científica-técnica con componentes de innovación para Reconstrucción del Estado de Juego y su traducción a métricas tácticas operativas.

1.3 Objetivos del proyecto

El objetivo principal es desarrollar un sistema de visión por computador end-to-end que detecte y siga jugadores, identifique dorsales, proyecte posiciones al campo calibrando la cámara y derive métricas y visualizaciones tácticas para apoyar decisiones futbolísticas informadas.

1.4 Resultados obtenidos

Se implementa un pipeline con YOLOv8+ByteTrack para la detección de jugadores, clustering por color, preprocesado cromático, OCR Tesseract para dorsales y calibración HRNet+PnLCalib con homografías. El resultado final son posiciones reales de los jugadores sobre las que se generaron diagramas de Voronoi, de control de campo continuo, de eventos SOG/SGG, de EPV y de ventaja sin balón con visualización 2D funcional en secuencias SoccerNet.

1.5 Estructura de la memoria

La memoria comprende: estado del arte, justificación, objetivos, metodología, resultados, discusión, conclusiones, y futuras líneas de trabajo.

Capítulo 2. ANTECEDENTES / ESTADO DEL ARTE

2.1 Estado del arte

El fútbol moderno se ha estado transformando radicalmente en los últimos años, evolucionando desde un deporte basado en intuición y observación hasta una disciplina que depende para muchos temas de los datos y análisis científico. La implementación de sistemas basados en inteligencia artificial para el análisis de partidos de fútbol ha representado una revolución en la forma en que entrenadores, analistas y equipos comprenden el juego, optimizan la toma de decisiones y mejoran el desempeño de los jugadores en el terreno de juego. El objetivo fundamental de estos sistemas es extraer información táctica y de rendimiento que permita una comprensión más profunda de la dinámica del juego, procesando cantidades masivas de datos recopilados durante entrenamientos y competiciones.

2.1.1 Desarrollo de la Inteligencia Artificial en el Deporte

En los últimos años, la integración de la inteligencia artificial en el análisis deportivo ha seguido una trayectoria marcada por muchos avances tecnológicos. Los primeros sistemas se basaban en métodos estadísticos básicos, pero estos han ido evolucionando hacia sistemas más sofisticados y personalizables de análisis de rendimiento deportivo [1]. Uno de los ejemplos más famosos es el sistema Hawk-Eye en tenis, introducido a principios de los 2000, que utiliza múltiples cámaras sincronizadas y algoritmos de triangulación para rastrear trayectorias de pelotas con una alta precisión [2] [3] [4].

En el contexto del fútbol, la aplicación de algoritmos de aprendizaje automático demuestra capacidades excepcionales para procesar información sobre la efectividad y precisión de las acciones de cada jugador. Los sistemas modernos identifican posibles posiciones basadas en

acciones más allá de la ubicación física, utilizando técnicas como Random Forest que han alcanzado precisiones del 64.78 % en predicción de resultados [5] [6] . Estos sistemas analizan una infinidad de características por partido, incorporando tanto acciones técnico-tácticas como variables contextuales [6] .

Los algoritmos de optimización también han pasado a tener un peso importante en la gestión deportiva, en temas como la valoración de jugadores y exploración de talentos. Modelos híbridos que combinan Decision Tree Regression con algoritmos metaheurísticos como Honey Badger Algorithm y Jellyfish Search Optimizer demuestran una precisión superior, alcanzando valores R^2 de 0.984 en predicción de valores de mercado. Estas implementaciones proporcionan alternativas confiables a estimaciones basadas en opiniones de expertos, lo cual pone en duda su importancia para el sector [7] .

Cabe destacar el caso de Scout Advisor del Sevilla FC, que utiliza modelos de lenguaje avanzados para la identificación de posibles talentos [8] [9] [10] . El sistema analiza miles de jugadores simultáneamente, identificando prospectos que coinciden con filosofías de juego específicas. También la introducción de la inteligencia artificial en la tecnología “wearable” ha cambiado en gran medida el monitoreo de la salud y rendimiento, recopilando datos continuos sobre estado físico, esfuerzo, recuperación y sueño [11] [12] [13] [14].

2.1.2 Evolución de la Visión por Computador, Arquitecturas Fundamentales y Rendimiento

La visión por computador se fundamenta en redes neuronales convolucionales que aplican filtros convolucionales para extraer características en diferentes niveles de abstracción. En estos momentos destacan las arquitecturas YOLOv7 y YOLOv8 para detección de objetos deportivos (aunque también son ampliamente utilizadas en otras industrias), las cuales muestran capacidades altas para detección simultánea de jugadores, árbitros y balones con precisiones superiores al 99%. La arquitectura que YOLO utiliza como backbone (CSPDarknet53) proporciona

balance óptimo entre precisión y eficiencia computacional, alcanzando precisiones del 97 % en DFL-Bundesliga y 98 % en SoccerNet-Tracking [15] [16] .

En estos tiempos los algoritmos de seguimiento como ByteTrack son populares debido a su consistencia en entornos dinámicos como campos de fútbol. Su algoritmo aborda el problema de mantener alta precisión en escenarios con múltiples objetos mediante re-asociación de objetos perdidos y mantenimiento de consistencia a través de varios fotogramas seguidos. Su formulación del seguimiento de múltiples objetos como problema de emparejamiento en grafos bipartitos demuestra mejoras considerables en métricas de intercambio de identidades y precisión de identificación como HOTA [17] [18] .

La convergencia de inteligencia artificial y visión por computador ha creado un ecosistema que transforma fundamentalmente el análisis deportivo. Ahora somos capaces de extraer y procesar datos visuales en tiempo real, permitiendo esto obtener información precisa y objetiva, registrando exhaustivamente todos los sucesos de un partido de cada jugador. Esta integración permite descubrir la forma de ser de los jugadores y evaluar la efectividad y precisión de sus acciones [15] . La detección de eventos clave durante competiciones representa también otra aplicación muy importante, ya que permite identificar de forma instantánea momentos cruciales y acelera la generación de resúmenes y estadísticas posteriores al partido para que los entrenadores y jugadores puedan revisarlos [18] .

2.1.3 Análisis Táctico en Fútbol: Metodologías y Aplicaciones

El análisis táctico en fútbol es el estudio exhaustivo de estrategias mediante observación y desglose de secuencias de partidos, estadísticas de jugadores y adquisición de datos en tiempo real. Este enfoque permite comprender comportamientos individuales y de equipo, despliegues estratégicos y dinámica del juego [19]. Los principios de juego son ideas fundamentales que forman un método de trabajo para preparar al jugador ante problemas tácticos y formar

mentalidad táctica, ya que es el jugador el primero que debe entender el porque de las decisiones tácticas que se le asignan.

Se ha observado que la investigación en análisis defensivo revela factores como presión defensiva, superioridad numérica cerca del balón y organización compacta, que son muy importantes a la hora de recuperar el balón. Los modelos de Random Forest han demostrado capacidad para identificar estos patrones tácticos con puntuaciones F1 muy buenas [20] . Mediante técnicas de agrupamiento jerárquico y algoritmos de Longest Common Subsequence, investigadores han identificado múltiples patrones frecuentes de velocidad y aceleración específicos para cinco roles principales: delantero, centrocampista, lateral, extremo y defensa central. Los sistemas de análisis de formaciones dinámicas utilizan estas técnicas de agrupamiento para computar formaciones de equipo, basándose en las posiciones de defensores y bloques de equipo. La aplicación de inteligencia artificial en el modelado de formaciones permite análisis en tiempo real que permiten tomar decisiones tácticas [21] .

También se han implementado herramientas tecnológicas especializadas como Az Sports que demuestran efectividad en minimización del margen de error en análisis táctico [22] . Los datos obtenidos permiten conocer los fuertes y carencias de equipos, generando una ventaja a la hora de generar estrategias. Los sistemas modernos pueden procesar información táctica con latencias menores a 100 milisegundos, proporcionando retroalimentación prácticamente instantánea sobre formaciones dinámicas y movimientos tácticos [23] .

2.1.4 Reconstrucción del Estado de Juego: Solución Integral

El problema de la Reconstrucción del Estado de Juego emerge como solución integral que unifica componentes especializados para generar representaciones comprehensivas del juego. Se define como la tarea de reconocer el estado completo mediante identificación y seguimiento de todos los atletas.

Para esto, el conjunto de datos SoccerNet-GSR ha proporcionado la primera referencia estándar, comprendiendo gran cantidad de secuencias de video segmentadas en frames completamente anotadas con más de 9.37 millones de puntos de línea para localización de campo y calibración de cámara, además de 2.36 millones de posiciones de atletas con información completa de rol, equipo y número de camiseta [24] . La métrica GS-HOTA específicamente desarrollada considera tanto precisión de seguimiento como correctitud de identificación, proporcionando evaluación más holística [24] .

El problema de la reconstrucción se subdividiría por tanto en tres problemas principales: localización de jugadores, reconocimiento de números de camiseta, y estimación de la posición relativa mediante calibración de cámara. Los métodos de detección de los jugadores en los distintos frames utilizan arquitecturas de aprendizaje profundo que predicen simultáneamente cajas delimitadoras y confianza de la predicción [25] . El enfoque Angle-Digit Refine Scheme combina información de ángulos de orientación corporal con claves de número de dígitos, demostrando precisiones de 64.07 % en Top-1 y 89.97 % en Top-2 [26] .

En calibración de cámaras se utilizan propiedades geométricas conocidas del campo como objeto de calibración. El pipeline PnLCalib combina detección de puntos clave con optimización de líneas para calibración robusta, utilizando redes de alta resolución como backbone para predicción de puntos clave y extremidades de líneas. La generación de puntos clave explota intersecciones línea-línea y línea-cónica, incrementando significativamente el número de puntos utilizables para calibración. Los sistemas más exitosos logran errores de reproyección menores a 0.005 metros e intersección sobre unión superiores a 98 % [27] .

2.1.5 Aplicación al Análisis Táctico Avanzado

La implementación exitosa permite hacer análisis de formaciones dinámicas e identificar patrones de transición previamente invisibles, permitiendo a los entrenadores visualizar en

tiempo real cómo equipos ocupan espacio e identificar oportunidades para ajustes tácticos. El sistema TacticAI, desarrollado en colaboración con Liverpool FC, permite explorar configuraciones alternativas de jugadores en saques de esquina. Los expertos demuestran preferir las sugerencias del modelo antes que las tácticas existentes en 90 % de los casos [28].

Además, los sistemas permiten cálculo automático de métricas sofisticadas como formaciones dinámicas de equipo, roles de jugadores, tipos de ataque y defensa. La disponibilidad de datos de posición precisos permite mejora significativa de modelos de valor esperado de posesión (EPV), proporcionando evaluaciones más exactas de creación de valor [29] . Por otro lado, los mapas de control del terreno modelan áreas de influencia basándose en posición, velocidad y capacidades físicas, permitiendo una evaluación del control territorial por parte de cada equipo. Otro modelo algo más antiguo es el de teselación de Voronoi, que permite ver patrones de ocupación espacial representando polígonos [30] .

También existen sistemas de análisis individualizado que permiten personalización de roles y responsabilidades tácticas basados en datos de rendimiento durante un partido. La aplicación de aprendizaje por refuerzo permite simular jugadas y sugerir movimientos óptimos basados en resultados previos [31] . Esto da pie a la existencia de entrenadores artificiales, que pueden crear ejercicios adaptados a debilidades específicas de jugadores, proporcionando hojas de ruta personalizadas para mejora.

2.1.6 Desafíos Tecnológicos y Direcciones Futuras

Una de las limitaciones que enfrentan los sistemas actuales es la transferencia de modelos entrenados en transmisiones profesionales a transmisiones de partidos amateur debido a diferencias en calidad de grabación y edición [32] . La dependencia en grandes cantidades de

datos anotados representa otra limitación, ya que la anotación manual es extremadamente costosa [33] .

La solución más prometedora en este aspecto es el desarrollo de métodos con supervisión mínima o aprendizaje auto-supervisado, con técnicas que aprovechen consistencia temporal y restricciones geométricas para reducir la dependencia en anotación manual [34] o la integración de fuentes multimodales, incluyendo audio, comentarios de texto y datos ambientales, que podría mejorar la robustez y la precisión de las predicciones [35] . Otra forma de mejorar los modelos existentes es mediante técnicas de optimización de modelos, incluyendo cuantización, poda y destilación de conocimiento [36] .

La democratización de análisis avanzados podría nivelar el campo entre clubes grandes y pequeños. Equipos que previamente no se podían permitir sistemas de seguimiento caros pueden ganar acceso a conocimientos similares a través de análisis de material de transmisión. El mercado global de análisis deportivo está creciendo significativamente, lo cual refleja la creciente importancia del análisis táctico basado en datos en el fútbol profesional [37].

2.2 Contexto y justificación

El fútbol constituye un deporte de dinámica compleja y alta variabilidad que ha experimentado una transformación significativa con la irrupción de la inteligencia artificial (IA) y el análisis de datos. Se ha demostrado que la capacidad de extraer información detallada del comportamiento de los jugadores, las tácticas de equipo y los patrones de juego en tiempo real ofrece un potencial sin precedentes para revolucionar la forma en que se entiende, entrena y disputa este deporte. Tradicionalmente, el análisis táctico y de rendimiento se ha basado en la observación humana y estadísticas descriptivas, métodos que son subjetivos, intensivos en tiempo y limitados en su capacidad para procesar la ingente cantidad de datos que se generan durante un partido.

La visión por computador se presenta como una tecnología disruptiva en este contexto, ya que, mediante el procesamiento automático de imágenes y secuencias de video de los partidos, permite obtener datos sobre el posicionamiento de los jugadores, sus movimientos, las interacciones con el balón y la configuración espacial del juego. Esta información, puede desvelar patrones ocultos, cuantificar el rendimiento de manera objetiva y proporcionar insights personalizados para entrenadores, analistas y los propios jugadores. Estos análisis ofrecen una perspectiva más completa y objetiva del rendimiento deportivo.

En este Trabajo de Fin de Máster (TFM) se desarrolla un sistema integral basado en visión por computador que se ha diseñado específicamente para analizar partidos de fútbol. El sistema propuesto abarca desde la detección e identificación de jugadores hasta el análisis de jugadas específicas, con el objetivo último de generar información que facilite la optimización de la toma de decisiones tácticas y la mejora del rendimiento individual y colectivo. A través de la aplicación de técnicas avanzadas de procesamiento de imágenes y modelos de aprendizaje profundo, se busca no solo automatizar tareas de análisis que actualmente se realizan de forma manual, sino también descubrir nuevas métricas y perspectivas que enriquezcan la comprensión del juego desde un enfoque más científico y cuantitativo.

La relevancia de este proyecto radica en su potencial para democratizar el acceso a herramientas de análisis avanzado, que hasta ahora han estado mayormente al alcance de clubes de élite con grandes recursos económicos y técnicos. El sistema no solo se plantea como una herramienta de análisis post-partido, sino que sienta las bases técnicas y metodológicas para futuras aplicaciones en tiempo real que puedan asistir en la toma de decisiones durante el transcurso del juego, proporcionando así un valor añadido significativo para el desarrollo del fútbol moderno.

2.3 Planteamiento del problema

Como se puede observar en el análisis del estado del arte, se ha producido una brecha crítica entre los avances técnicos individuales y su integración práctica en sistemas de análisis táctico

futbolístico. Aunque se han desarrollado tecnologías como YOLOv8 que alcanzan precisiones superiores al 99 % en detección de jugadores y algoritmos como ByteTrack que mantienen seguimiento robusto, la fragmentación tecnológica impide que se generen soluciones comprehensivas y accesibles. Los sistemas actuales se han especializado en tareas específicas que operan como silos independientes: detección de jugadores, reconocimiento de números de camiseta, calibración de cámara y análisis táctico funcionan por separado, lo que dificulta considerablemente su integración coherente. El conjunto SoccerNet-GSR representa un avance hacia la unificación mediante la Reconstrucción del Estado de Juego, pero se requiere la coordinación de múltiples componentes especializados que utilizan diferentes marcos de referencia y formatos de datos. Paralelamente, se ha identificado una brecha de accesibilidad pronunciada entre clubes de élite que disponen de sistemas como Scout Advisor del Sevilla FC o TacticAI desarrollado con Liverpool FC, y las organizaciones deportivas con recursos limitados, lo que contradice el potencial democratizador que se podría alcanzar mediante la visión por computador aplicada a material de transmisión estándar.

Se ha evidenciado una desconexión significativa entre las capacidades técnicas alcanzadas por los sistemas de visión por computador y su aplicación práctica en análisis táctico. Aunque se logra detectar y seguir jugadores con alta precisión técnica, la traducción de datos posicionales en insights tácticos verdaderamente significativos permanece subdesarrollada en la mayoría de las implementaciones actuales. Los sistemas existentes se limitan a generar mapas de calor básicos y métricas elementales de posicionamiento, pero carecen de las capacidades necesarias para identificar patrones tácticos complejos, evaluar la efectividad de transiciones defensiva-ofensiva en tiempo real, o proporcionar recomendaciones estratégicas contextualizadas que resulten verdaderamente accesibles para cuerpos técnicos. La ausencia de frameworks unificados para análisis táctico integral obliga a que entrenadores y analistas dependan de interpretación manual de información fragmentada, lo que limita significativamente la eficiencia del proceso analítico e introduce elementos de subjetividad que contrarrestan los beneficios de objetividad que se buscan mediante la automatización.

La problemática descrita revela la necesidad de desarrollar un sistema integral que se enfrente simultáneamente a las limitaciones identificadas. Se requiere una solución que unifique los

componentes especializados existentes en un pipeline coherente y robusto que procese automáticamente desde la detección inicial hasta la generación de insights tácticos, que democratice el acceso mediante una implementación computacionalmente eficiente compatible con hardware estándar y material de transmisión convencional, que proporcione robustez operacional minimizando la dependencia en anotación manual costosa, y que genere análisis táctico mediante la traducción inteligente de datos posicionales en recomendaciones estratégicas específicas y contextualizadas.

El desarrollo de esta solución integral se plantea como una contribución tanto técnica al campo de la visión por computador aplicada al deporte como un avance sustancial hacia la democratización efectiva del análisis táctico avanzado, con potencial transformador para que equipos de todos los niveles puedan comprender y optimizar su rendimiento deportivo de manera objetiva y científica.

Capítulo 3. OBJETIVOS

3.1 Objetivos generales

El presente Trabajo de Fin de Máster tiene como objetivo explorar y aplicar tecnologías de inteligencia artificial, en particular la visión por computador, para desarrollar e implementar un sistema que analice de forma avanzada partidos de fútbol a partir de secuencias de vídeo, capaz de interpretar y extraer información táctica y de rendimiento que contribuya a una comprensión más profunda del juego, optimice la toma de decisiones y la estrategia, mejore el desempeño de los jugadores en el terreno de juego y transforme datos visuales brutos en conocimiento accionable para entrenadores, analistas y jugadores.

3.2 Objetivos específicos

Para la consecución del objetivo general, se han definido los siguientes objetivos específicos que guiarán el desarrollo del proyecto:

- Desarrollar e implementar un modelo robusto para la identificación, selección y seguimiento de jugadores en secuencias de vídeo de partidos de fútbol. Esto implicará la aplicación de algoritmos de detección de objetos (como YOLO o similares) para localizar a los jugadores en cada fotograma. También se desarrollarán mecanismos para asignar una identidad única a cada jugador detectado y mantener su seguimiento a lo largo de diferentes fotogramas, gestionando occlusiones y cambios de apariencia.
- Establecer un flujo de trabajo eficiente para el procesamiento de imágenes individuales para su posterior análisis. Se investigarán técnicas para la extracción óptima de fotogramas, considerando la frecuencia de muestreo necesaria para capturar los

eventos relevantes del juego sin incurrir en una sobrecarga computacional innecesaria.

Se gestionarán los datos de entrada y salida (secuencias de imágenes y metadatos asociados) de manera organizada.

- Extraer y analizar análisis relevantes a partir de la información posicional y de movimiento de los jugadores, identificando jugadas y momentos clave del partido. Se buscará identificar eventos como la formación de líneas defensivas, la creación y ocupación de espacios, los patrones de movimiento en ataque y defensa, y la posesión del balón. Se explorarán métodos para la detección automática de eventos significativos a partir de las trayectorias y configuraciones de los jugadores.
- Elaborar un sistema capaz de generar informes o representaciones visuales que sugieran posibles mejoras en jugadas específicas o aspectos tácticos del equipo. Basándose en los insights obtenidos, se diseñarán mecanismos para visualizar y comunicar información útil, como la identificación de espacios libres no explotados, la evaluación de líneas de pase óptimas, o la sugerencia de movimientos alternativos en situaciones de juego específicas. Se explorará la posibilidad de cuantificar la efectividad de ciertas jugadas o la contribución de los jugadores a la creación de oportunidades.

3.3 Beneficios del proyecto

El desarrollo e implementación del sistema propuesto genera múltiples beneficios tangibles que responden directamente a los objetivos planteados y abordan las necesidades identificadas en el análisis del estado del arte.

Uno de los beneficios principales, como venimos diciendo, es la democratización del análisis táctico avanzado. Mediante el desarrollo de un sistema que opera con material de transmisión

estándar y hardware accesible, el trabajo elimina las barreras económicas y técnicas que tradicionalmente han restringido el acceso a herramientas de análisis sofisticado únicamente a clubes de élite. Este beneficio responde directamente al objetivo principal de extraer información táctica y de rendimiento, pero lo hace de manera inclusiva, permitiendo que equipos de diferentes niveles competitivos puedan acceder a informes que anteriormente requerían inversiones considerables en sistemas especializados como los utilizados por el Sevilla FC o Liverpool FC.

Otro beneficio fundamental es la automatización integral del pipeline de análisis. La implementación de un flujo de trabajo automatizado desde la segmentación de video hasta la generación de informes visuales elimina la dependencia en análisis manual intensivo en tiempo y recursos humanos. Esta automatización no solo incrementa la eficiencia operativa, sino que también reduce la subjetividad inherente en métodos tradicionales de análisis, proporcionando evaluaciones más objetivas y consistentes del rendimiento deportivo. El sistema permite procesar partidos completos de manera autónoma, generando datos posicionales precisos y métricas tácticas sin requerir intervención humana constante.

Un beneficio directo para los usuarios finales del sistema es la generación de análisis tácticos profundos, ya que el proyecto no se limita a replicar métricas básicas como mapas de calor o distancias recorridas, sino que desarrolla capacidades analíticas avanzadas que identifican patrones de formación dinámica, espacios no explotados, oportunidades de creación de juego y evaluación cuantitativa de efectividad táctica. Estos insights permiten a entrenadores y analistas tomar decisiones informadas basadas en evidencia cuantitativa, optimizando estrategias tanto a nivel individual como colectivo.

Por último, la escalabilidad y adaptabilidad del sistema desarrollado permite su aplicación en diversos contextos futbolísticos, desde competiciones amateur hasta ligas profesionales. Esta característica responde al objetivo de crear una herramienta funcional, pero sobrepasa sus limitaciones iniciales, proporcionando una plataforma que puede adaptarse a diferentes calidades de video, formatos de transmisión y niveles de competición. La arquitectura modular

Angel Molpeceres Mingo

del sistema facilita mejoras futuras y personalización según necesidades específicas de diferentes organizaciones deportivas.

Capítulo 4. DESARROLLO DEL PROYECTO

4.1 Planificación del proyecto

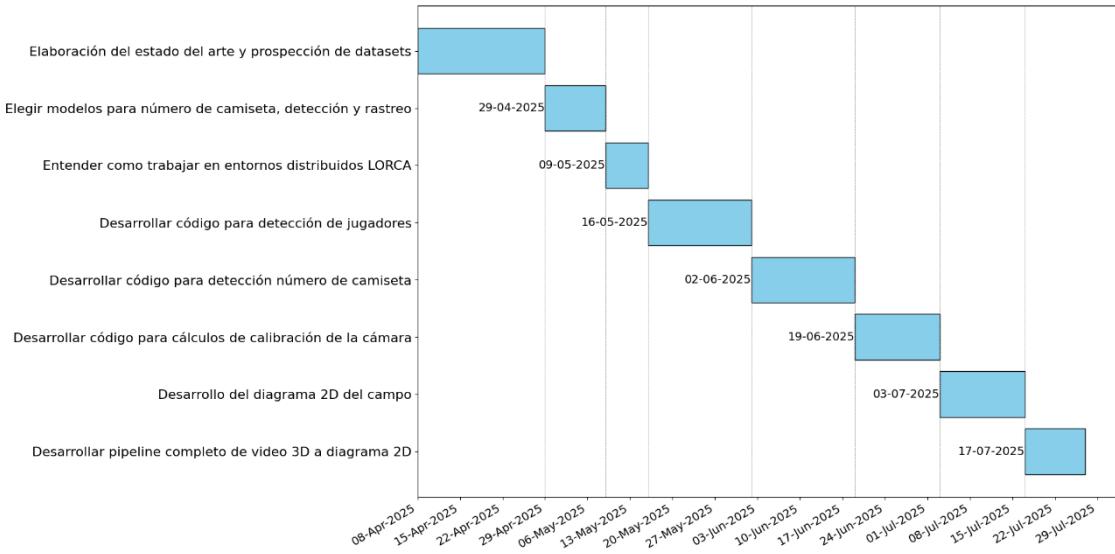


Figura 1. Diagrama de Gantt de la planificación del proyecto

En un principio se diseñó de esta manera la planificación y división en partes del proyecto, reservando una parte para el trabajo en entornos distribuidos usando el centro de procesamiento de datos de la Universidad Europea LORCA, pero se acabó desestimando por una cuestión de gestión de tiempos.

4.2 Descripción de la solución, metodologías y herramientas empleadas

4.2.1 Adquisición y preparación de los datos

El proceso se inició con la obtención del kit SoccerNet-Tracking-2023, que incluye secuencias de 15 partidos, utilizando la API oficial de SoccerNet a través de Python.

Se implementó un sistema que descarga tanto los datos del desafío de seguimiento original como los específicos del desafío 2023, proporcionando acceso a diferentes versiones del conjunto de datos para análisis comparativo. SoccerNet Downloader permite la descarga automática de datos y anotaciones mediante un paquete pip especializado. El paquete descargado contiene conjuntos referentes a dos tareas principales.

Por un lado, se encuentra la tarea de "tracking", correspondiente al conjunto de datos original de SoccerNet-Tracking, que incluye 200 secuencias de 30 segundos cada una, representativas de escenarios desafiantes de fútbol, además de un tiempo completo de 45 minutos completamente anotado para seguimiento a largo plazo. Por otro lado, se presenta la tarea "tracking-2023", una versión actualizada que se centra en el desafío completo de seguimiento que requiere detectar los objetos de interés desde el vídeo en bruto, a diferencia del desafío 2022 que consideraba detecciones ground-truth.

Los datos descargados siguen una estructura jerárquica bien definida que facilita su procesamiento y análisis. El conjunto de datos está separado en carpetas para train y test. Dentro de estas carpetas se encuentran secuencias con el nombre SNMOT-XXX que incluyen una carpeta llamada gt con el archivo txt de ground truth, una carpeta llamada img1 con los 750 frames de cada secuencia y otra serie de archivos secundarios de distinta índole. El conjunto de datos utiliza el formato estándar MOT Challenge para las anotaciones almacenadas en los archivos de ground truth, que contienen 10 columnas. La estructura de cada línea de anotación sigue el formato de la Tabla 2.

Frame ID	Track ID	X	Y	W	H	Score	-1	-1	-1
1	1	914	855	55	172	1	-1	-1	-1
2	1	907	855	67	172	1	-1	-1	-1

Tabla 2. Formato de anotaciones en archivos de ground truth según el estándar MOT Challenge.

Donde Frame ID es el identificador del frame en la secuencia, Track ID es el identificador único del objeto rastreado, X, Y son las coordenadas de la esquina superior izquierda de la bounding box, W, H son el ancho y alto del bounding box, Score es la puntuación de confianza (1 para ground truth) y los últimos valores no utilizados marcados como -1 según los requerimientos del formato MOT20. Esta estructura permite implementar algoritmos de seguimiento que procesan secuencialmente cada imagen, manteniendo la coherencia temporal necesaria para el seguimiento de múltiples objetos.

Durante la realización del proyecto nos basaremos en esta estructura para nuestros archivos de detección de jugadores, que incluirán no solo las coordenadas de las bounding boxes sino también números de jugador, pertenencia a un equipo u otro, etc.

4.2.2 Desarrollo del pipeline de procesamiento

La detección de jugadores en cada frame constituye el primer paso fundamental de nuestro pipeline de procesamiento, para esta tarea se implementa YOLOv8, que representa la evolución de la familia YOLO (You Only Look Once), implementando un enfoque de detección en una sola pasada que predice simultáneamente coordenadas de bounding boxes y probabilidades de clase. El algoritmo cuenta con una primera fase en la que la imagen de entrada es procesada por una profunda red convolucional basada en CSPDarknet53, que analiza la imagen y extrae un conjunto de características visuales muy amplio en múltiples escalas (tanto detalles finos como patrones más abstractos). El número de canales de salida se calcula como:

$$C_{out} = \min (C_{base} \times w_{mult}, C_{max}) \times d_{mult}$$

donde w_{mult} es el multiplicador de ancho, d_{mult} el multiplicador de profundidad y C_{max} el límite máximo de canales [38] .

Las características extraídas se integran en una segunda fase usando el mecanismo PANet, que combina la información semántica de alto nivel (los que ayudan a identificar lo que es el objeto) con los detalles de bajo nivel (que definen el lugar en el que está el objeto). Finalmente, esta información llega a los módulos de salida, separados en módulos de clasificación que asignan una etiqueta al objeto y módulos de regresión que definen el tamaño y ubicación de la bounding box. Estos últimos utilizan una función de pérdida llamada Distribution Focal Loss, que optimiza el cálculo de las coordenadas de la caja y tiene la siguiente fórmula:

$$DFL(p_i) = \sum_{j=0}^n (S_j \log(p_{i+j}) + S_{j+1} \log(p_{i+j+1}))$$

donde S_j es la etiqueta suavizada y p_{i+j} la probabilidad predicha para el bin j [39].

Este modelo fue seleccionado por su capacidad para detectar múltiples objetos simultáneamente en tiempo real, característica esencial para el seguimiento de jugadores en entornos dinámicos como un partido de fútbol.

El proceso de detección se inicia con la carga secuencial de cada frame del vídeo, manteniendo la estructura de datos proporcionada por SoccerNet. Se configura el modelo YOLOv8 con un umbral de confianza de 0.75 para minimizar los falsos positivos, asegurando que solo las detecciones con alta probabilidad sean consideradas para el procesamiento posterior. Esta configuración es crucial para distinguir eficazmente entre jugadores, árbitros y otros elementos presentes en el campo. Este segmento del pipeline nos permite trazar las “bounding boxes” o recuadros que rodean la figura de cada jugador en el campo, obteniendo así las coordenadas de estos en cada imagen. De estos recuadros se guardan las coordenadas de la esquina superior izquierda y la altura y ancho de los mismos.

Antes de proceder con el resto del pipeline, se debe hacer un paso más, que concierne a los id's de las detecciones de cada jugador en cada frame, número que se les da previo a la detección del número real de su camiseta. Uno de los paradigmas principales en la detección de recuadros de los jugadores es la permanencia de la pertenencia de una bounding box de un jugador de un frame a otro. Para esto se aplica ByteTrack un framework que aporta una solución al problema de asociación de detecciones siguiendo un proceso de asociación dual, donde primero se realizan detecciones de alta confianza $score \geq \tau_{high}$ las cuales se asocian con tracklets activos mediante distancia IoU:

$$IoU(b_i, b_j) = \frac{|b_i \cap b_j|}{|b_i \cup b_j|}$$

Donde $|b_i \cap b_j|$ es el área de la intersección y $|b_i \cup b_j|$ el área de la unión entre las dos cajas delimitadoras [40].

A continuación, se hace una segunda asociación de detecciones de baja confianza $score \leq \tau_{high}$ se asocian con tracklets perdidos usando un umbral más permisivo. El estado de cada tracklet se modela mediante un filtro de Kalman con el siguiente vector de estado:

$$x = [u, v, s, r, \dot{u}, \dot{v}, \dot{s}]^T$$

Donde u, v son las coordenadas centrales de recuadro, s su área, r su relación de aspecto y $\dot{u}, \dot{v}, \dot{s}$ sus derivadas temporales [41].

Esto permite una cierta consistencia de los id's para cuando la diferencia de posición de los recuadros de un mismo jugador en dos frames es reducida y no hay situaciones de movimiento repentino de la cámara, superposición de varias bounding boxes u oclusión del jugador.

Una vez detectados los jugadores, se implementa un algoritmo de clustering basado en K-means para clasificarlos según el color de sus equipaciones. Este enfoque permite diferenciar automáticamente entre los dos equipos, porteros y árbitros sin necesidad de etiquetado manual. Primero se extraen las características de color en el espacio HSV (Hue Saturation Value) para cada jugador detectado, después se aplica el algoritmo de clustering K-means con K=3 para los tres grupos principales, equipo local, equipo visitante y un tercer grupo de porteros y árbitros. Por último, se asignan etiquetas de clase basadas en los centroides de color resultantes. Esta clasificación por color es fundamental para el seguimiento posterior, ya que permite mantener la identidad de los jugadores incluso en situaciones de oclusión parcial o cambios de iluminación.

Previo a poder leer los números de las camisetas, primero se debe preparar las imágenes. Para ello, se utiliza el recuadro (*bounding box*) que detecta a cada jugador en un fotograma para crear un recorte individual. A continuación, estos recortes se someten a un preprocesado para mejorar su calidad y facilitar el reconocimiento de los dígitos. El flujo de trabajo para preparar los recortes de jugadores consta de tres subetapas principales: recorte central del fotograma, mejora de resolución y realce selectivo de color. En primer lugar, se eliminan el 20 % superior e inferior de cada fotograma para concentrar el análisis en el torso y dorsales del jugador. Este recorte elimina elementos irrelevantes (gradas, cabeza, piernas) y reduce el ruido de fondo. El procedimiento toma la altura original del fotograma, calcula los límites de recorte en función del porcentaje especificado y extrae la sección central manteniendo el ancho completo del fotograma.

Una vez recortados los fotogramas, se aplica un método de superresolución sencillo basado en interpolación bicúbica. Este enfoque duplica las dimensiones de la imagen originales, suavizando transiciones y preservando detalles de bordes críticos para la forma de los números de dorsal. Esta interpolación utiliza una fórmula polinómica de tercer grado para calcular nuevos píxeles a partir de una vecindad de 16 píxeles originales, mejorando la nitidez sin introducir error significativo que sigue la siguiente fórmula:

$$f(x, y) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij} x^i y^j$$

Donde los coeficientes a_{ij} se calculan utilizando los 16 píxeles vecinos y sus derivadas.

Para dar mayor contraste a los dorsales y eliminar variaciones de iluminación, se combinan dos métodos de variación de color. El primer método es CLAHE (Contrast Limited Adaptive Histogram Equalization). En lugar de modificar los colores directamente, la imagen se convierte primero al espacio de color LAB, que separa la información de luminosidad (el canal L) de los componentes de color (los canales A y B). Esta separación permite ajustar el brillo y el contraste de la imagen sin alterar los colores originales de la camiseta o el fondo. El algoritmo CLAHE opera exclusivamente sobre este canal de luminosidad, dividiendo la imagen en pequeñas regiones y mejorando el contraste localmente en cada una de ellas. Además, evita que el ruido se amplifique en áreas de color uniforme, como el tejido de la camiseta, logrando así que los detalles del dorsal resalten tanto en las zonas más oscuras como en las más iluminadas de manera mucho más clara y definida. El segundo método consiste en una saturación selectiva en HSV. Se detectan áreas con alta componente de color (no grises) y se aumenta la saturación exclusivamente en esas regiones, enfatizando tonos de camiseta o números frente al césped y el público. La mezcla ponderada de ambas mejoras (más énfasis en saturación sobre contraste) asegura un realce equilibrado de dorsales y equipaciones sin deformar la apariencia del jugador.

Para simplificar la identificación de equipaciones y centrar la atención en los dorsales, se reduce la paleta de colores a ocho tonos básicos, eliminando así variaciones de iluminación y facilitando la separación visual de jugadores y fondo. Por un lado, los píxeles de la imagen mejorada se agrupan en ocho clusters en el espacio RGB mediante K-Means, obteniendo colores representativos de cada región visual. Después, cada centroide se reemplaza por el color básico más próximo de un conjunto predefinido (blanco, negro, rojo, verde césped, etc.), calculando la distancia euclídea en RGB y priorizando detección de verdes intensos para el césped. Una vez simplificados los colores de los recortes de cada jugador se procede a realizar un muestreo de

los colores más representativos según su proximidad al centro del recorte. Para ello, se calcula la distancia euclídea de cada píxel al centro geométrico y se seleccionan los primeros cuatro colores distintos detectados en orden de cercanía. De este proceso se obtienen cuatro nuevas imágenes en blanco y negro a partir de cada recorte jugador donde, en cada caso, el color seleccionado se cambiará por negro y el resto de los colores de la imagen se cambiarán a blanco para destacar únicamente ese color frente al resto.

Una vez aislados estos cuatro componentes de color en imágenes binarias, se aplica un análisis de componentes conectados con conectividad de ocho vecinos para identificar las regiones potenciales que corresponden a dígitos. Se validan los grupos detectados según criterios de área máxima relativa (40 % del total), no superposición excesiva en columnas verticales (80 % de altura) y ausencia de contacto con los bordes de la imagen, priorizando las dos regiones más grandes y con alturas similares para cubrir dorsales de uno o dos dígitos. Antes de la etapa de OCR, cada región binarizada de los dígitos seleccionados es sometida a un redimensionamiento proporcional con interpolación cúbica, ajustando la dimensión mayor a 600 px para equilibrar detalle y tiempo de procesamiento. Esta normalización mejora la legibilidad de los caracteres sin introducir artefactos que puedan confundir al motor OCR.

Para la extracción de los dígitos, se utiliza el motor OCR Tesseract sobre cada componente binarizado y las predicciones se integran a lo largo de varios fotogramas mediante un esquema de votación por mayoría, descartando instancias con baja confianza. Tesseract es un pipeline de varias etapas que procesa la imagen desde componentes conectados hasta el reconocimiento del carácter final. Primeramente, realiza un análisis de componentes conectados inspeccionando la jerarquía de contornos, lo cual permite el reconocimiento de texto negro sobre fondo blanco. Después se usan splines cuadráticos para modelar líneas base curvadas:

$$y(x) = ax^2 + bx + c$$

Los parámetros se estiman mediante regresión por mínimos cuadrados sobre los centroides de los caracteres detectados. Por último, se hace una clasificación de caracteres empleando una métrica de distancia simétrica entre características del carácter desconocido:

$$d(x, \mu_k) = \frac{1}{M} \sum_{l,i} |x_{il} - \mu_{ijk}|^2 + \frac{1}{N} \sum_{j,i} |x_{il} - \mu_{ijk}|^2$$

donde x representa el carácter desconocido y μ_k el prototipo con el que se compara. Las variables x_{il} y μ_{ijk} son las características individuales de cada uno (como segmentos de línea o bucles), mientras que M y N son el número total de características que posee el carácter desconocido y el prototipo, respectivamente [42] .

Para robustecer el reconocimiento, se implementa una validación temporal que consolida las predicciones de los dígitos detectados en ventanas de varios frames. De esta forma, se mitigan efectos de occlusiones parciales o deformaciones del dorsal debidas a movimientos rápidos, asegurando continuidad en la identificación y reduciendo los cambios erráticos de número a lo largo de la secuencia. Tras la ejecución de esta parte del pipeline se obtendría un archivo en formato txt de detecciones de jugadores con la información obtenida por los algoritmos que tiene la siguiente estructura de la Tabla 3.

Frame ID	Track ID	X	Y	W	H	Score	Team ID	DetectedNum
1	1	285.00	490.00	80.00	149.00	0.8423 91	Local	['15']
1	2	1171.00	397.00	57.00	108.00	0.9021 90	Visitante	[]

Tabla 3. Estructura de las detecciones de jugadores por frame

Tenemos para cada frame, una serie de detecciones de los distintos jugadores, con una línea para cada una de estas detecciones, con coordenadas, altura y ancho de la bounding box de cada uno, un score de confianza de la detección (a partir de 0.75), el equipo al que pertenece y el numero detectado para el jugador en ese frame concreto. Una vez identificados los jugadores y sus características en el plano de la imagen, el desafío principal consiste en traducir sus coordenadas de píxeles a una representación métrica y objetiva sobre el terreno de juego. Este proceso de transformación de perspectiva es fundamental para el análisis táctico y se basa en una calibración precisa de la cámara en cada fotograma, para lo cual se aprovecha la geometría del propio campo de fútbol como referencia.

El primer paso para lograrlo es el análisis de la imagen mediante redes neuronales profundas. Específicamente, se utilizan dos modelos con una arquitectura High-Resolution Network (HRNet), un modelo está entrenado para detectar las líneas del campo (banda, fondo, áreas, círculo central), mientras que el otro se especializa en identificar puntos de interés (intersecciones, esquinas, punto de penalti). HRNet ha sido elegida por su capacidad para mantener representaciones de alta resolución a lo largo de todo el proceso de inferencia mediante conexiones paralelas multi-resolución. Su innovación clave radica en el intercambio repetido de información entre subredes de diferentes resoluciones.

Primeramente, incorpora una función de fusión multi-escala que permite que la información de todas las ramas de resolución (alta, media, baja) se intercambie y fusionen en cada una de las etapas de esta función. Para ello, las características de una de las ramas de resolución se escalan para coincidir con las demás y luego se suman, permitiendo que el contexto semántico de las ramas de baja resolución mejore la precisión espacial de las ramas de alta resolución. Para una subred de resolución r en el paso p :

$$Y_p^r = \sum_{r'}^{N_p} f_{r',r}^p X_p^{r'}$$

Donde $f_{r',r}^p$ es la función de transformación que resuelve r' a r , implementada mediante convoluciones 1x1 y operaciones de escalado.

Después, en lugar de predecir coordenadas directamente, la red genera un "mapa de calor" para cada punto clave, que funciona como un mapa de probabilidad. Para el entrenamiento, la ubicación real del punto clave se representa como el centro de una distribución Gaussiana 2D. La red aprende a replicar este mapa, y la predicción final se obtiene localizando el punto de máxima activación en el mapa de calor generado. La salida final genera mapas de calor para cada punto clave k :

$$H_k(x, y) = \exp\left(-\frac{(x - x_k)^2 + (y - y_k)^2}{2\sigma^2}\right)$$

donde (x_k, y_k) son las coordenadas reales del punto clave y σ controla la dispersión de la distribución Gaussiana [43].

Para este trabajo se eligen puntos clave como esquinas y puntos medios de los lados del área, los círculos, la intersección de la línea de medio con la banda, etc. El resultado de estos modelos no son coordenadas directas, sino mapas de calor: imágenes en escala de grises donde la intensidad de cada píxel representa la probabilidad de que forme parte de una línea o un punto clave. A partir de estos mapas de calor, un algoritmo de postprocesamiento extrae las coordenadas 2D de los puntos y líneas con mayor probabilidad, que sirven como entrada para el framework de calibración PnLCalib. Este componente es clave, ya que implementa una solución robusta al problema clásico de visión por computador conocido como Perspective-n-Point (PnP). Dicho método utiliza las correspondencias entre los puntos 2D detectados en la imagen y su ubicación conocida en un modelo 3D ideal del campo de fútbol para estimar los parámetros extrínsecos de la cámara: su rotación y traslación en el espacio. De este modo, se obtiene una comprensión tridimensional completa de la perspectiva de la cámara en ese instante.

Si bien la calibración PnP proporciona una orientación 3D completa, para la tarea específica de mapear jugadores sobre el césped, se calcula una matriz de homografía mediante la siguiente fórmula:

$$s \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = H \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

donde el punto (x,y) en el plano del campo se proyecta como el punto (x', y') en el plano de la imagen mediante la matriz de homografía H . El factor de escala s se utiliza para manejar las coordenadas homogéneas [44].

Esta es una transformación matemática 2D-a-2D que relaciona directamente el plano de la imagen con el plano del campo. Se calcula utilizando los pares de puntos de correspondencia (píxeles en la imagen y metros en el campo). Para garantizar la máxima fiabilidad, el cálculo se realiza mediante el algoritmo RANSAC (Random Sample Consensus), que descarta eficazmente las detecciones atípicas o erróneas que podrían distorsionar el resultado, un factor crítico en secuencias con movimiento rápido u occlusiones. Con la matriz de homografía lista para cada fotograma, se procede a mapear a los jugadores. Para cada detección, se define como punto de anclaje el punto central del borde inferior de la bounding box, ya que esta ubicación se corresponde con el contacto de los pies del jugador con el césped. Al aplicar la transformación de homografía a este punto, sus coordenadas en píxeles se convierten en coordenadas (x, y) reales y en metros dentro del plano del campo.

El resultado de este complejo pipeline se consolida en un archivo txt final que unifica toda la información relevante. La estructura de este archivo está diseñada para ser fácilmente interpretable y utilizada en fases posteriores y se muestra en la Tabla 4.

Frame ID	Track ID	X_Real	Y_Real	Team ID	DetectedNum
1	1	64.07	61.47	Local	['2']
1	2	40.89	29.05	Visitante	['13']

Tabla 4. Coordenadas reales de los jugadores en el plano del campo tras la transformación por homografía

Estos datos son los que finalmente alimentan la generación de la vista táctica 2D, donde cada jugador es representado como un punto en un diagrama del campo a escala, permitiendo un análisis visual y objetivo de las formaciones, movimientos y posicionamiento táctico a lo largo del partido.

4.2.3 Generación de insights de partido

Una vez mapeadas con precisión las coordenadas de los jugadores sobre el plano del campo, se dispone de una base de datos que sirve como punto de partida para un análisis táctico de nivel superior. En esta fase, se aplican una serie de modelos analíticos para transformar estos datos brutos en métricas y visualizaciones que revelan la dinámica espacial del juego, el control del terreno y el valor de las acciones de los jugadores.

Para generar una primera capa de análisis territorial, se implementa el cálculo de diagramas de Voronoi a través de la biblioteca `scipy.spatial`. Los diagramas de Voronoi dividen el campo en regiones poligonales asignadas a cada jugador, de modo que cada punto pertenezca al jugador más cercano. Esta partición permite conocer qué jugador llegaría antes al balón si este quedase libre. La simplicidad de este método lo hace ideal para una primera aproximación al control territorial, aunque asume velocidades uniformes y no contempla la dinámica de movimiento. En la Figura 2 se puede observar un ejemplo de diagrama de Voronoi.

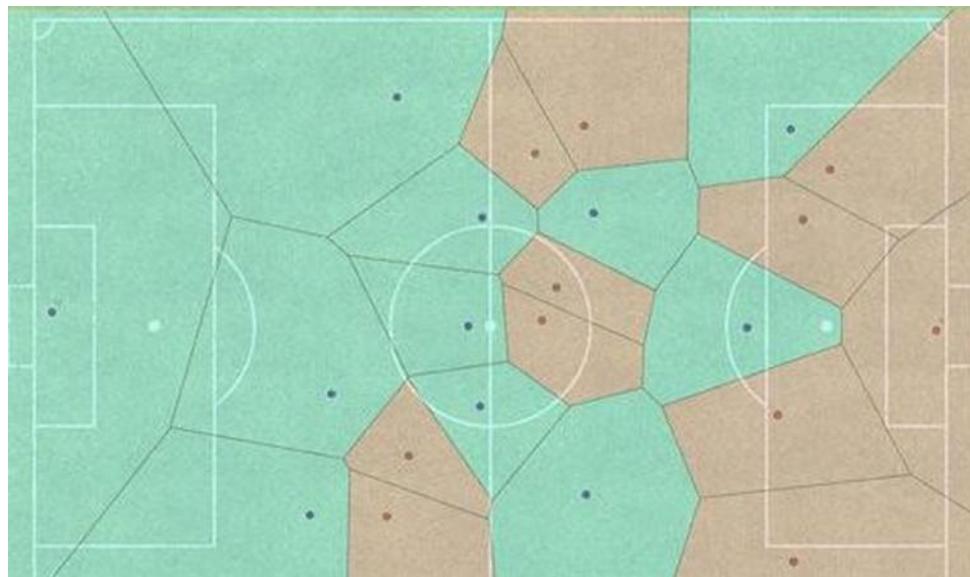


Figura 2. Generación de insights tácticos mediante diagramas de Voronoi para el análisis del control espacial del terreno de juego [49].

Para resolver el problema de las regiones infinitas que se producen en los bordes del campo, se añaden puntos auxiliares o "fantasma" a una distancia considerable fuera de los límites del terreno de juego. Esta técnica fuerza al algoritmo a generar polígonos cerrados para todos los jugadores, asegurando que cada uno tenga un área de influencia definida. Finalmente, se colorea cada región de Voronoi según el equipo al que pertenece el jugador y se superpone el resultado con un cierto nivel de transparencia sobre la vista táctica 2D. De esta forma, se ofrece un mapa claro del dominio espacial inmediato de cada equipo.

Con el objetivo de obtener una visión más sofisticada del dominio espacial, se implementa un modelo de control del campo el cual trasciende la rigidez de Voronoi al tratar la influencia de cada jugador como un campo continuo y decreciente. Cada futbolista proyecta una “colina de influencia” modelada por una distribución gaussiana bidimensional:

$$f_i(p) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{|p - \mu^i|^2}{2\sigma^2}\right)$$

Donde μ^i es la posición media, desplazada ligeramente hacia la dirección de movimiento, y σ controla el alcance.

Para cada punto p del campo, se suman las influencias de los jugadores de ambos equipos, $PC_A(p)$ y $PC_B(p)$, y su diferencia normalizada mediante una tangente hiperbólica:

$$PC(p) = \tanh(PC_A - PC_B)$$

El resultado es un mapa de calor continuo con valores entre -1 y $+1$, que indican la probabilidad de control del balón por parte de cada equipo [46]. Podemos observar en la Figura 3 un ejemplo de mapa de control del campo.

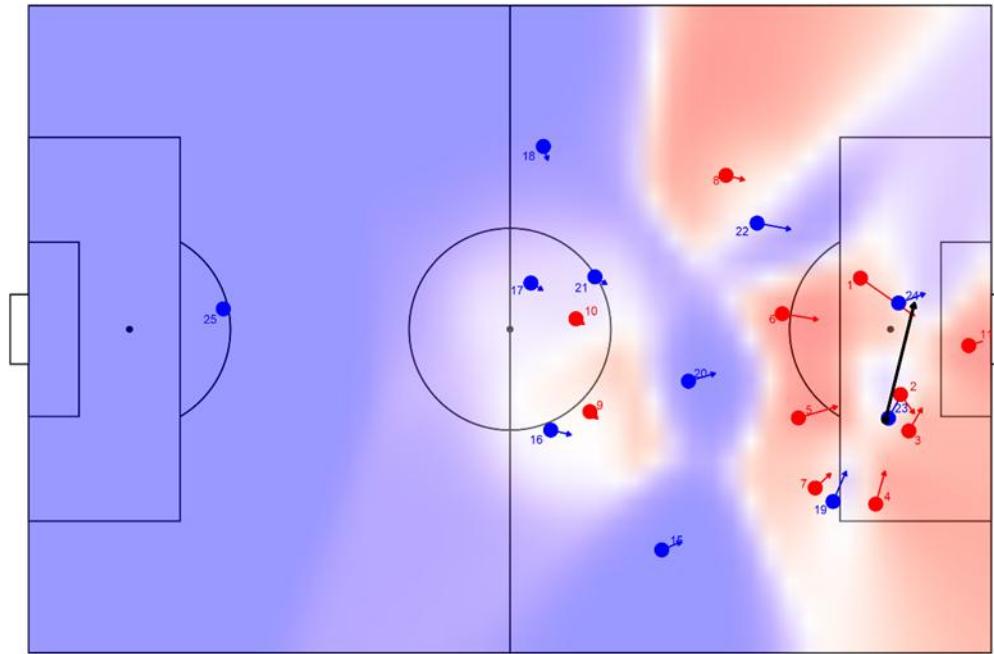


Figura 3. Mapa de control del campo [50] .

Sobre la base del control del campo, se implementan también la detección de eventos de creación y ocupación de espacio. El primer paso es definir un mapa de valor espacial, que

consiste en un gradiente lineal que asigna un valor creciente desde la propia portería hasta la rival.

$$V(x, y) = \frac{x}{L}$$

Donde x es la distancia al área rival y L la longitud total del campo.

Se calcula la calidad de espacio Q de un jugador i multiplicando el valor de control del campo de su equipo en su ubicación p^i por el valor espacial correspondiente a ese punto.

$$Q^i = PC(p^i) \times V(p^i)$$

Se define un evento de ocupación de espacio (SOG) como el momento en el que el jugador mejora significativamente su calidad de espacio en un breve intervalo:

$$\Delta Q = Q_t^i - Q_{t-\delta}^i > \theta_{SOG}$$

Con un umbral típico para θ_{SOG} de 0.05. Si la velocidad de desplazamiento excede de 7km/h, se clasifica como SOG activo; de lo contrario, pasivo. Un evento de generación de espacio (SGG) identifica cuándo un “generador” atrae a un defensor, liberando a un “receptor” que antes estaba marcado y ahora queda en espacio libre [47]. En la Figura 4 se puede observar este fenómeno.



Figura 4. Secuencia de generación de espacio. Elaboración propia.

En este trabajo, la detección de eventos se realiza comparando el estado actual con un estado anterior, que se almacena en un búfer histórico de 10 fotogramas:

- Evento SOG (Ganancia de Ocupación de Espacio): Se registra este evento cuando la calidad de espacio de un jugador aumenta por encima de un umbral predefinido. Adicionalmente, se clasifica el evento como "activo" o "pasivo" según si la velocidad del jugador en ese intervalo supera los 7 km/h. Visualmente, se destaca a los jugadores que realizan un SOG con un aura (amarilla para activo, azul para pasivo).
- Evento SGG (Generación de Espacio para un Compañero): Se implementa una lógica que identifica una interacción entre tres jugadores: un "generador" que arrastra a un "defensor" (al acercarse a él) y, como consecuencia, libera a un "receptor" (que se aleja del mismo defensor). Se visualiza estos eventos con una línea discontinua que conecta al generador y al receptor.

Se desarrolla un modelo de EPV para evaluar en tiempo real el peligro potencial de una jugada. Este modelo teóricamente combina los factores: posición y superioridad numérica. La posición (VP) es la proximidad del portador a la portería rival normalizada como:

$$VP = \frac{x_{portador}}{L}$$

La Superioridad numérica (SN) es, por su lado, la proporción de atacantes y defensores por delante del balón, ajustada a un rango [0.5, 1.5], se obtiene por ponderación lineal:

$$EVP = 0.7VP + 0.3(SN - 0.5)$$

Un incremento brusco de EPV ($\Delta EPV > 0.15$) señala una carrera de valor, y se atribuye al jugador con mayor avance [48] .

En el presente trabajo, el valor de EPV se muestra en la interfaz como una barra de estado. A partir de esta métrica, se implementa la detección de carreras de valor, que se registran cuando el EPV aumenta bruscamente por encima de un umbral en el intervalo de tiempo analizado. El sistema identifica al jugador atacante con el mayor desplazamiento hacia adelante como el principal responsable y resalta su trayectoria con un rastro de color en la vista 2D. Para simplificar el cómputo se incorpora un parámetro que establece que el valor del espacio siempre aumenta de izquierda a derecha. Esto implica que el equipo que ataca hacia la derecha es considerado el equipo atacante y se puede cambiar cuando el caso es el contrario.

Finalmente, se implementa un modelo para identificar a los receptores de pase más ventajosos en una jugada de ataque, conocido como modelo de cálculo de ventaja sin balón. Esta métrica resalta a los jugadores mejor posicionados para recibir un pase sin oposición inmediata. Se definen dos condiciones: posicional $\frac{x_i}{L} > 0.6$, es decir, en el último tercio del campo y presión, a lo sumo un defensor en un radio de 5 m. Cuando se cumplen ambas, el jugador se considera en ventaja, señalando “bolsas de espacio” explotables en el ataque [47].

Se marca a los jugadores que cumplen ambos criterios en la visualización 2D con un aura blanca parpadeante, señalando así al portador del balón cuáles son las opciones de pase más claras y peligrosas para progresar en el ataque.

4.3 Recursos requeridos

Para la realización del proyecto se ha utilizado un portátil HP con un procesador Intel Core I5 de 11^a generación y con 8 GB de RAM. Se ha elaborado código en Python para la resolución de las distintas partes de la metodología.

4.4 Presupuesto

Tipo de coste	Valor	Comentarios
Horas de trabajo en el proyecto	250 horas	Repartidas principalmente durante los meses de mayo, junio y agosto
Equipo técnico utilizado	600 €	Ordenador HP
Software utilizado	0 €	VSCode, Python 3.12.5
Estudios e informes	0 €	Estudios e informes disponibles de forma pública
Sueldo técnico	3750 €	Sueldo de Data Scientist Junior (30.000 €/año)

Tabla 5. Presupuesto del proyecto

4.5 Viabilidad

El sistema desarrollado presenta una viabilidad económica sólida debido a su arquitectura basada en tecnologías open-source que minimiza los costes de licenciamiento y desarrollo. La

inversión consistiría en el hardware de procesamiento (500-600 €) y el tiempo de desarrollo ya invertido, mientras que los costes operacionales serían 0,05 € por minuto de vídeo procesado. Esta estructura permite implementar modelos de negocio escalables, desde suscripciones SaaS de 50-200 € mensuales para clubes amateur a contratos de licenciamiento de 10,000-50,000 € anuales para federaciones o plataformas deportivas establecidas.

El mercado objetivo abarca a los más de 10,000 clubes federados en España, con el segmento internacional de la analítica deportiva experimentando un crecimiento anual del 22%. La propuesta de valor se centra en democratizar herramientas de análisis táctico que tradicionalmente solo estaban disponibles para clubes de élite, aprovechando la brecha existente entre sistemas como TacticAI del Liverpool FC y las necesidades analíticas de equipos con recursos limitados. El uso de material de transmisión estándar elimina la dependencia de infraestructuras costosas, ampliando significativamente el mercado accesible.

4.6 Resultados del proyecto

4.6.1 Resultados del pipeline

El pipeline desarrollado se aplica a las distintas secuencias de partido de 750 frames disponibles, que incluyen distintos tipos de jugadas, desde faltas y córneres hasta penaltis y contraataques.



Figura 5. Aplicación del pipeline a secuencias de partido: imagen de partido

En la Figura 5 podemos observar uno de los muchos frames pertenecientes a las secuencias con las que se trabaja.

En la Figura 6 se observa el primer paso del proceso, que consiste en la detección de bounding boxes alrededor de los jugadores en pantalla. Al reducir la confianza mínima a 0.75, tal como se indica en la metodología, se detectan como jugadores las siluetas de personas dentro del campo, lo que permite minimizar las detecciones fuera del campo, entre las que podrían incluirse espectadores del partido o personal de seguridad.

El resultado obtenido se muestra a continuación en la Figura 6.



Figura 6. Detección de jugadores en el campo con asignación de bounding boxes

Se observa que todos los jugadores tienen asignada su correspondiente bounding box. También se visualiza un identificador que corresponde con el ID asignado a cada detección, el cual mantiene cierta permanencia al cambiar de frame, siempre que no exista una superposición significativa entre uno o más jugadores. En caso de superposición, podría perderse la traza.

Para evitar depender de este identificador y, además, conocer la identidad real de los jugadores, se aplica la siguiente parte del pipeline: el reconocimiento de números.



Figura 7. Reconocimiento de números de dorsal en jugadores detectados.

En el frame de la Figura 7 se identifican unos números de dorsal visibles. Se analiza cómo actúa el sistema de reconocimiento de números sobre estos recortes en particular. En primer lugar, se selecciona la fracción superior de la imagen, lo que facilita el enfoque en el número y reduce el ruido visual (Figura 8).



Figura 8. Recorte y procesamiento de la región del dorsal.

A continuación, se aplica un filtro de superresolución que mejora la nitidez de la imagen y permite separar con mayor claridad los números del contorno en el que se encuentran. Este paso se muestra en la Figura 9, aunque a simple vista no se pueda detectar.



Figura 9. Mejora de nitidez mediante superresolución.

La mejora de la nitidez visualizada en la Figura 9, permite una mayor separación visual entre los números de dorsal y el césped, lo cual resulta especialmente relevante en situaciones donde el contraste es reducido. A continuación, se realiza un ajuste en los colores de la imagen (Figura 10), incrementando la saturación, mejorando el contraste y aplicando una serie de transformaciones adicionales.



Figura 10. Transformación cromática y reducción de complejidad visual.

Seguidamente se reduce la complejidad cromática de la imagen, limitando la paleta a unos pocos colores básicos. Esta transformación (que se puede observar en la Figura 11) permite obtener nuevos recortes con una representación más simplificada, facilitando el análisis posterior.



Figura 11. Reducción de la paleta cromática.

Tal como se ha indicado previamente, el objetivo es aislar el color del número respecto a los colores que lo rodean, mejorando así su identificación. Este efecto se intensifica en el siguiente paso visualizado en la Figura 12, en el que se limita la imagen a una paleta de únicamente ocho colores.



Figura 12. Simplificación cromática mediante paleta limitada a ocho colores.

En el siguiente paso mostrado en la Figura 13 se seleccionan los cuatro colores con mayor presencia en la imagen. Cada uno de estos colores se aísla secuencialmente, asignándole el color negro, mientras que el resto de los colores de la imagen se transforman en blanco. Esta

operación permite obtener una serie de recortes con alto contraste, adecuados para el análisis de formas y patrones.



Figura 13. Aislamiento secuencial de los colores dominantes.

Con un ejemplo representativo por cada color es suficiente para ilustrar el procedimiento de segmentación cromática. En el caso de la equipación blanca (el jugador número 28), se procede a aislar únicamente tres colores, en lugar de cuatro, debido a la menor complejidad cromática del uniforme.

El siguiente paso del pipeline incorpora una lógica más avanzada. Su objetivo principal es identificar y extraer exclusivamente los grupos de píxeles negros que representan dígitos correspondientes a los números de dorsal. Para ello, se emplea un algoritmo de detección de componentes conectados, que identifica agrupaciones de píxeles negros y las filtra en función de criterios predefinidos: tamaño máximo permitido, exclusión de componentes que toquen los bordes de la imagen, y eliminación de aquellos que ocupen columnas completas.

El sistema selecciona hasta dos componentes válidos que se encuentren a una altura similar, lo que permite gestionar correctamente dorsales compuestos por dos dígitos. Se incorpora un recuadro visual en la Figura 14 para facilitar la interpretación del contenido relevante en la imagen.

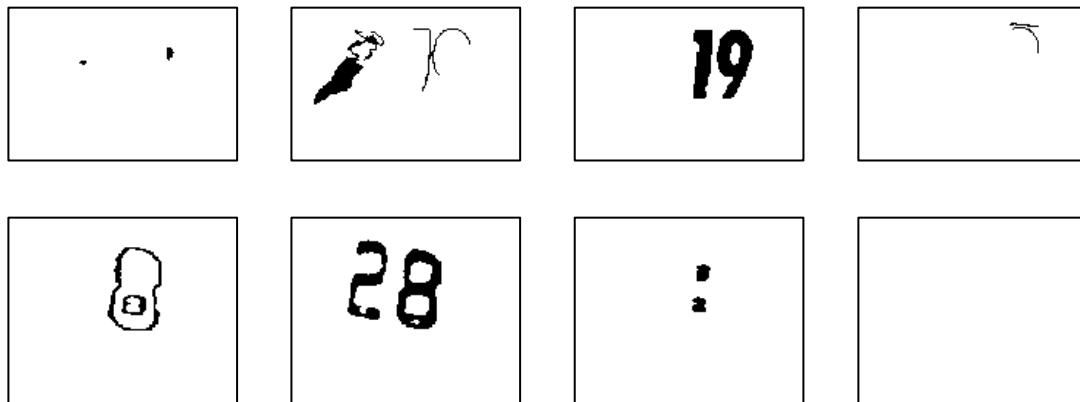


Figura 14. Aislamiento de dígitos en dorsales mediante selección de componentes visuales relevantes.

Los dígitos han sido correctamente aislados del ruido visual circundante, lo que permite avanzar a la fase final del pipeline: el reconocimiento de números.

Dado que el proceso se desarrolla de forma no supervisada, se aplica un sistema de reconocimiento óptico de caracteres (OCR) sobre todas las imágenes generadas. El análisis se realiza considerando todos los dígitos reconocidos en las cuatro imágenes correspondientes a cada jugador, lo que permite mejorar la fiabilidad del resultado mediante redundancia informativa. Se puede ver el resultado del código en la figura 15.

```
NumRecogFolder > resultados_ocr_minimo.txt
 1  RESULTADOS OCR CON PROCESAMIENTO MINIMO
 2  =====
 3
 4  SUBCARPETA: player_frame51_id2
 5  | componente_1.png: No detectado
 6  | componente_2.png: No detectado
 7  | componente_3.png: 19
 8  | componente_4.png: No detectado
 9
10  SUBCARPETA: player_frame51_id3
11  | componente_1.png: No detectado
12  | componente_2.png: 28
13  | componente_3.png: No detectado
14  | componente_4.png: No detectado
15
```

Figura 15. Reconocimiento óptico de caracteres (OCR) para la identificación final del dorsal.

En los casos analizados, los dígitos son claramente visibles y no se detectan componentes adicionales que interfieran en la segmentación. La secuencia presentada constituye una visualización detallada de los pasos que conforman el pipeline de reconocimiento de números. No obstante, este proceso no es observable directamente durante la ejecución, ya que únicamente se dispone de los resultados finales obtenidos en los frames tras el procesamiento.



Figura 16. Resultados completos del pipeline de reconocimiento de jugadores y dorsales.

Los resultados de la Figura 16 muestran una detección precisa de los números de dorsal en la mayoría de los casos. La única excepción se puede observar en la figura 17 y corresponde al dorsal 22, en el que únicamente se ha detectado el dígito 2, lo que indica una pérdida parcial de información durante el proceso de reconocimiento. Estas pérdidas son frecuentes y poco relevantes ya que se hacen detecciones de los números en el resto de frames y se detectaría el número correcto más adelante y en un mayor número de veces.



Figura 17. Excepción en el reconocimiento del dorsal.

El siguiente paso del pipeline consiste en inferir la localización real de cada jugador en el terreno de juego, a partir de su detección en cada frame de la secuencia. Para ello, se procede inicialmente a la identificación precisa de una serie de puntos clave en el campo. Conociendo la posición real de estos puntos en el terreno de juego, es posible calcular la ubicación espacial de los jugadores mediante técnicas de transformación geométrica.

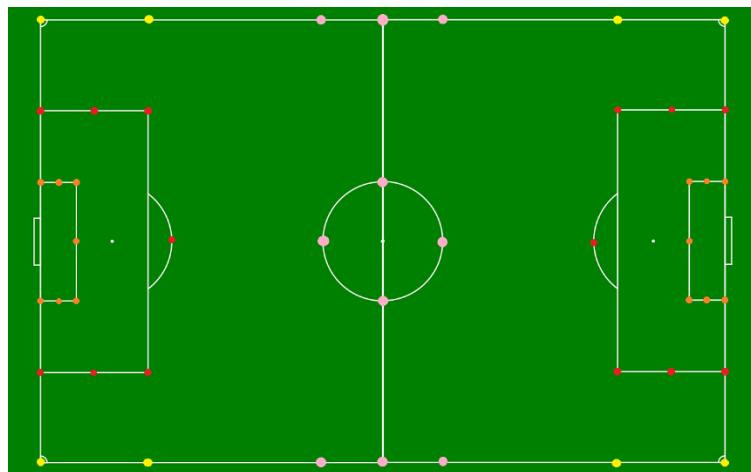


Figura 18. Identificación de puntos clave en el terreno de juego para la estimación geométrica de la posición real de los jugadores [51]

Se seleccionan los puntos que se pueden observar en la Figura 18 como referencias espaciales debido a su presencia constante en la mayoría de los planos de las secuencias, lo que garantiza

una base estable para la estimación de posiciones. Luego, se implementa un módulo específico para la detección automática de puntos en el terreno de juego, el cual permite identificar con precisión las marcas reglamentarias. Además, se muestra en la Figura 19 el funcionamiento del sistema sobre las líneas del campo, lo que facilita la validación visual del proceso de localización.

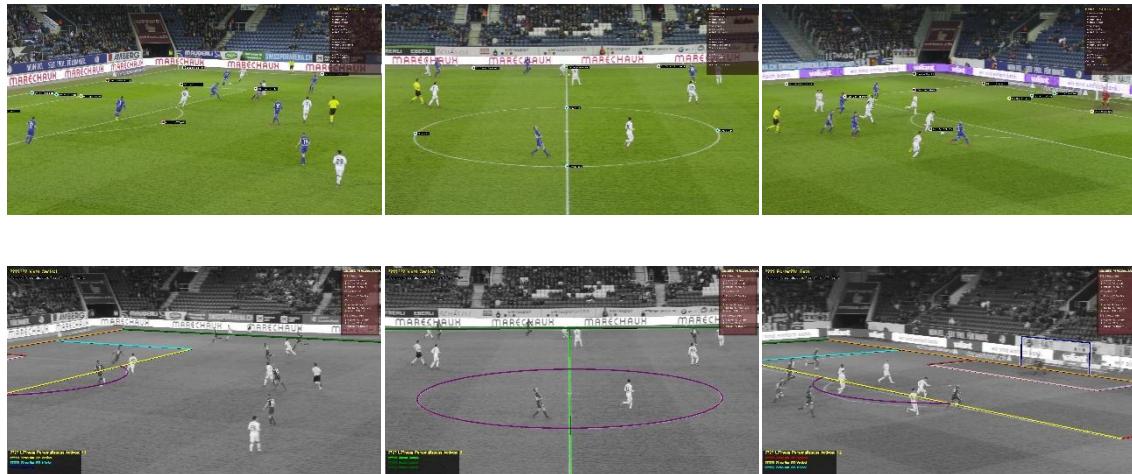


Figura 19. Detección automática de líneas del campo para validar la localización espacial de los jugadores.

A partir de los puntos clave detectados en cada frame y de la localización de los jugadores en cada instante, se calcula la matriz de homografía que permite proyectar las posiciones desde el plano de la imagen al plano del terreno de juego. El resultado se representa en un diagrama bidimensional del campo, donde se visualiza la distribución espacial de los jugadores en coordenadas reales (Figura 20).

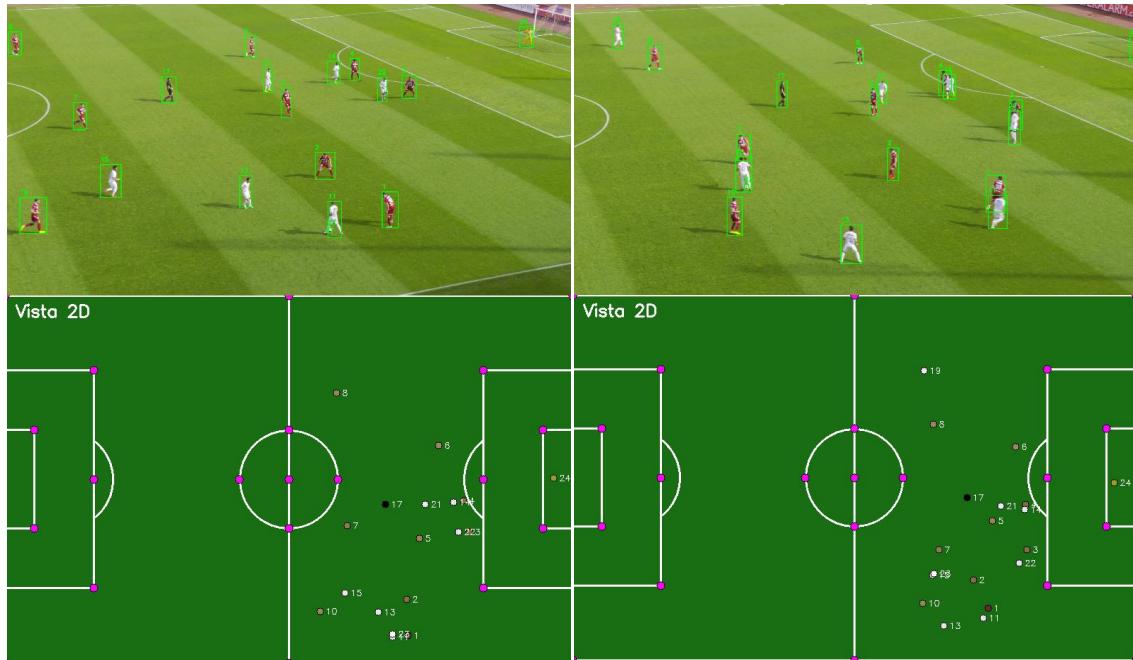


Figura 20. Proyección de posiciones reales de los jugadores sobre el plano en dos dimensiones del campo mediante homografía.

De esta forma, es posible visualizar secuencias completas del juego en un diagrama bidimensional del terreno. Este representa el resultado final del pipeline de reconstrucción, que abarca desde la detección y localización de los agentes en cada frame, pasando por el reconocimiento de los dorsales visibles, hasta la proyección de sus posiciones al sistema de coordenadas reales del campo.

4.6.2 Análisis de jugadas

En este apartado se analiza cómo aplicar los modelos de comprensión y análisis de toma de decisiones, descritos previamente en la metodología, a las distintas secuencias de jugadas disponibles.

Tal como se explicó anteriormente, los diagramas de Voronoi ofrecen una primera aproximación al análisis del espacio ocupado en el terreno de juego. Estos permiten evaluar, a posteriori, si los jugadores se encontraban bien posicionados en función del contexto táctico. Por ejemplo, en situaciones ofensivas, los delanteros deben estar abiertos a un pase profundo sin incurrir en fuera de juego. Este tipo de posicionamiento puede identificarse en los diagramas cuando el polígono de Voronoi de un jugador se extiende hasta los límites del campo, lo que indica que podría alcanzar un balón largo antes que los defensores.

En la secuencia analizada, se observa cómo el jugador con ID 22 se encuentra inicialmente en una posición abierta para recibir un pase (Figura 21 a) y, posteriormente, rota su posición para volver a quedar desmarcado y disponible para un pase en profundidad (Figura 21 c))

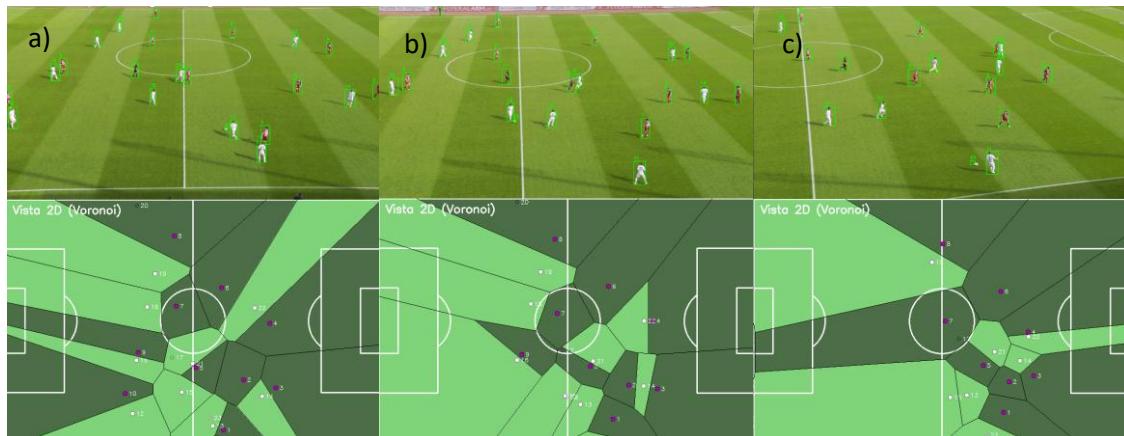


Figura 21. Evolución posicional del jugador ID 22 en secuencia táctica: apertura, rotación y desmarque.

De esta misma secuencia pueden identificarse momentos críticos en los que la pérdida de posesión representa un mayor riesgo para el equipo blanco. En la Figura 21 a), por ejemplo, se observa que los jugadores con ID 10 y 9 se encuentran en una posición muy favorable para iniciar un contraataque, siempre que el jugador 10 se retrase ligeramente para evitar el fuera de juego. Esta situación se refleja en la extensión de sus polígonos de Voronoi, que alcanzan los bordes del campo, indicando que podrían anticiparse a los defensores en la recepción de un pase largo. En cambio, en la Figura 21 c) no puede realizarse el mismo análisis, ya que no se dispone de la

localización completa de los defensores del equipo blanco, lo que impide evaluar con precisión el riesgo de pérdida en esa fase de la jugada.

Desde la perspectiva defensiva, es fundamental que ningún delantero al que se esté marcando disponga de un polígono de Voronoi amplio, como los descritos anteriormente, ya que esto indicaría una ventaja espacial significativa que podría traducirse en una oportunidad de gol.

El siguiente modelo, el de control de campo, que es un poco más complejo se puede observar las zonas controladas por un equipo u otro de una forma más realista y homogénea. El diagrama generado tiene un equipo rojo y otro azul, el balón es un círculo morado, no se representa de forma principal a los árbitros y los porteros no crean áreas de control de campo.

En la Figura 22 se puede ver como en una jugada de ataque del equipo rojo, el equipo azul se centra en controlar la zona del campo cercana al balón. Se observa también que un jugador del equipo rojo se abre a la otra banda, efectivamente controlando esa zona del campo, sin influencia del equipo azul.

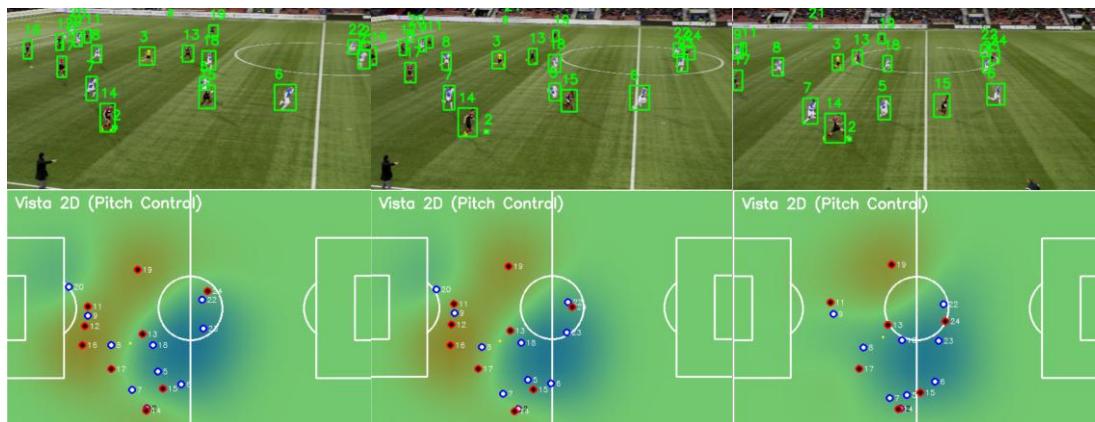


Figura 22. Distribución del control territorial durante una jugada ofensiva: concentración defensiva y apertura estratégica.

En esta otra secuencia que se observa en la Figura 23, se puede ver como durante el desarrollo de una jugada, van evolucionando las zonas de control de cada equipo, estando al principio muy distribuidas para los dos equipos, para después cada equipo repartir su espacio de control en zonas más concretas.

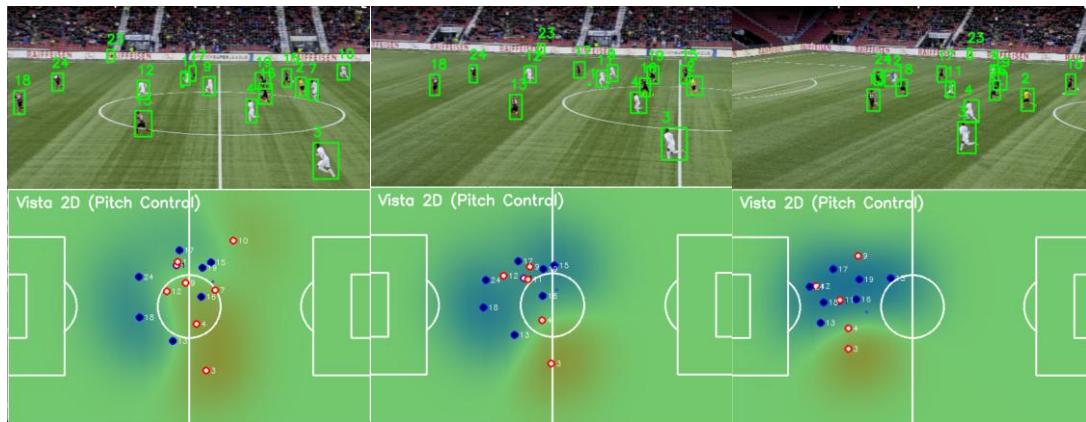


Figura 23. Evolución de las zonas de control durante una jugada: transición de distribución amplia a concentración táctica.

En esta tercera secuencia que se observa en la Figura 24 se puede ver una jugada en la que ambos equipos se disputan el balón en una zona que controla el equipo azul. Esto hace que un mal pase o fallo en el ataque del equipo rojo, acabe con la devolución de la posesión al equipo azul. Se observa además que tras recuperar el equipo azul la posesión, son capaces de capitalizar esa recuperación y hacer un pase a un jugador desmarcado, que se puede ver que se mete con el balón detrás incluso de la zona de control de los defensas del equipo rojo.

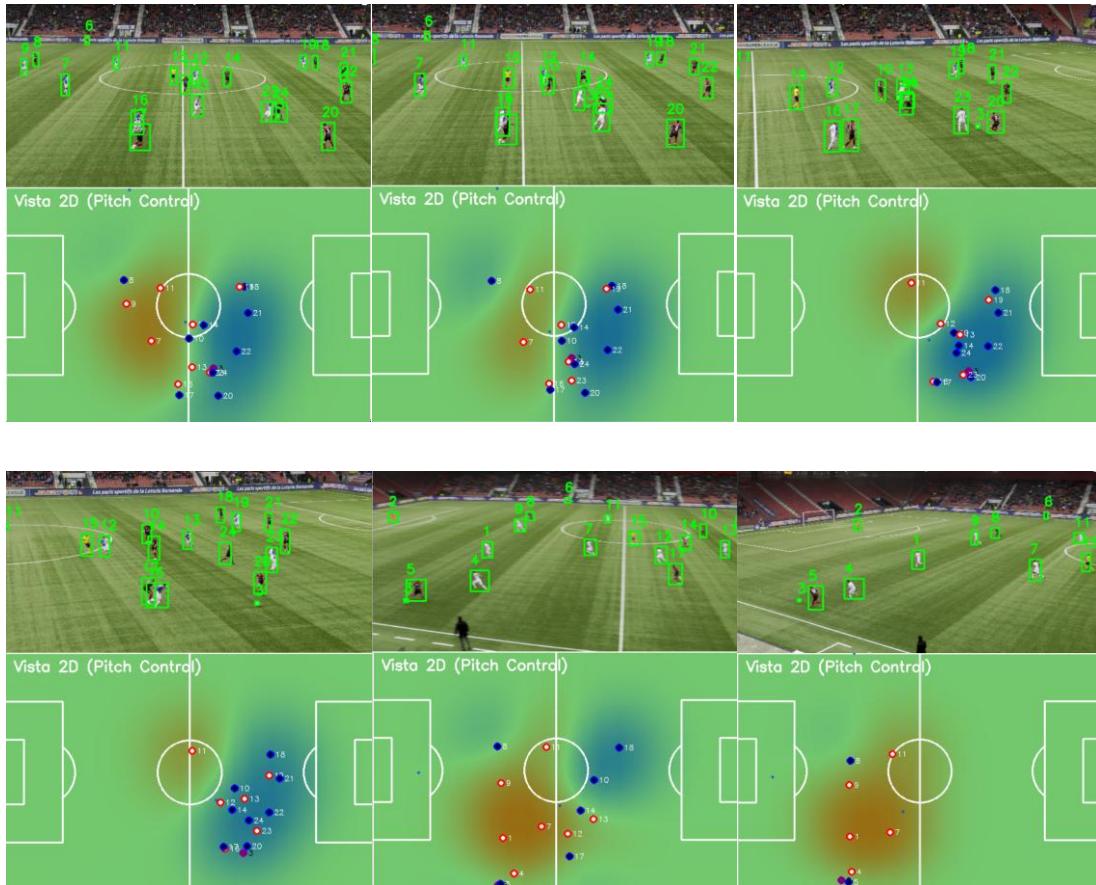


Figura 24. Disputa de posesión y aprovechamiento del espacio tras recuperación por parte del equipo azul.

En el siguiente modelo de visualización táctico de la Figura 25 se obtendría la identificación de eventos SOG y SGG. Se puede visualizar eventos SOG de ganancia de ocupación de espacios activos (aura amarilla) y pasivos (aura azul) en la siguiente secuencia de la Figura 24, en la que se puede ver el inicio de un contraataque del equipo negro, donde el jugador 13 realiza una carrera a un espacio vacío para asentar el control de su equipo de ese espacio.



Figura 25. Visualización de eventos SOG activos y pasivos durante el inicio de un contraataque del equipo negro.

De esta forma se podrían identificar movimientos clave de jugadores que permiten ampliar la porción del campo que controla cada equipo.

El siguiente evento que se podría detectar es el de generación de espacio para un compañero (SGG), en el que interactúa un generador del evento, el cual se llevaría a un defensor consigo haciendo que su compañero se quedase libre para recibir un balón. La secuencia que muestra la Figura 26 se observa una jugada en la que el equipo rojo tiene la posesión. Se observa al jugador con id 1 con el balón, marcado por el jugador 11 y cerca de él está su compañero con id 2 y un jugador contrario con id 13.



Figura 26. Generación de espacio mediante atracción defensiva y desmarque en jugada ofensiva del equipo rojo.

Se puede observar como el jugador 1 se trae al jugador 13 dejando al 2 libre sin cubrir para un pase, señalándolo el diagrama con una línea discontinua entre los dos jugadores del equipo atacante involucrados.

El siguiente modelo es el de valor esperado de la posesión (EPV) el cuál se implementa para detectar carreras de valor de los jugadores atacantes y que las resalta en el diagrama del mapa con una trayectoria de su desplazamiento detrás de él. En el caso de la Figura 27 se puede ver al equipo negro atacando y se observa cómo en el display se pueden ver las carreras valiosas para el equipo atacante, como podrían ser las carreras de los jugadores 15 y 14 al espacio libre en la banda para crear la jugada o el desmarque final del jugador 24 que acaba recibiendo un pase en profundidad y marcando gol.

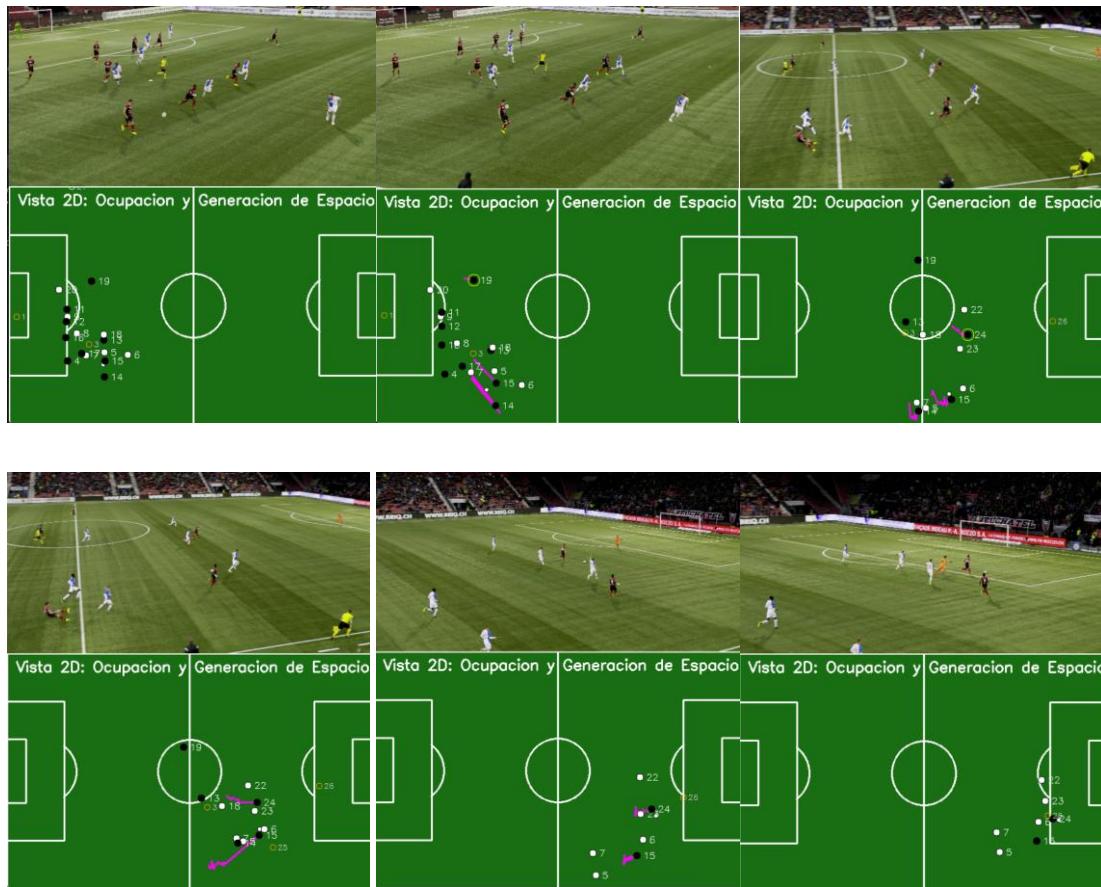


Figura 27. Visualización de carreras de valor en jugada ofensiva mediante modelo EPV: generación de oportunidades y finalización.

El último modelo es el de detección de ventaja sin balón, que identifica los jugadores del equipo atacante en una posición más ventajosa para un pase. En la Figura 28 se observa una jugada de ataque por parte del equipo blanco en la que al jugador 11, que lidera el ataque con el balón, se le presentan tres opciones prometedoras de pase.

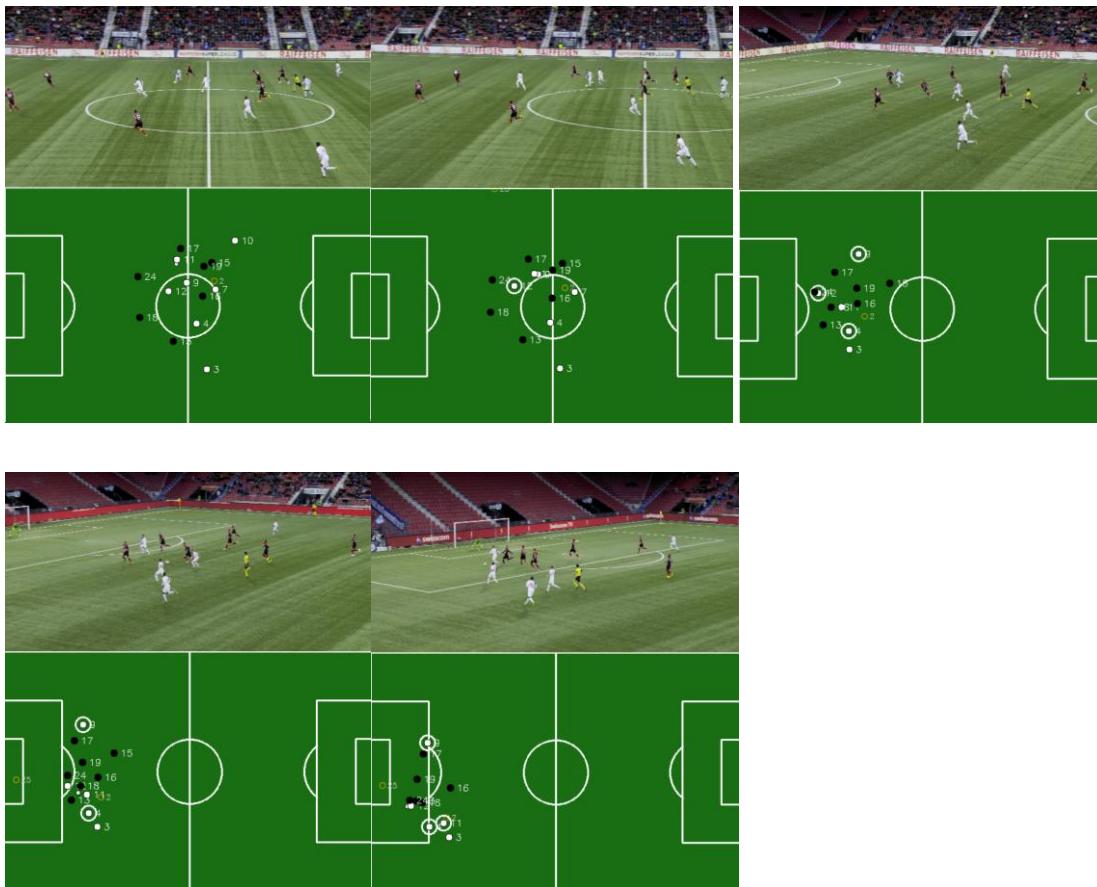


Figura 28. Identificación de opciones de pase mediante modelo de ventaja sin balón en jugada ofensiva del equipo blanco.

Como se ve, acaba optando por el jugador más adelantado, lo cual acaba ocasionando una ocasión de gol para su equipo.

Capítulo 5. DISCUSIÓN

Como se puede ver, el proyecto desarrollado constituye una respuesta al problema de la reconstrucción del estado de juego, de la que se pueden obtener unos resultados visuales útiles y completos que permitirían a clubes de fútbol obtener insights clave de partido. Todos los involucrados en el juego (entrenador, delanteros, defensas y centrocampistas) pueden obtener un informe de las acciones de todos los jugadores en el campo, ver tanto jugadas individuales como de equipo y aprender de sus errores.

Se ha optado por no incluir los informes de uso más antiguos y extendidos como son los mapas de calor de cada jugador (que muestran las zonas del campo que más a frecuentado) o, por ejemplo, los informes de velocidad media o distancia recorrida de cada jugador. Estos informes pecan de simples y en este trabajo se buscaba análisis más profundos que involucrasen conceptos matemáticos y aproximaciones al entendimiento del fútbol desde un punto de vista más científico.

La metodología propuesta ha permitido la implementación ordenada paso a paso de la solución desde la localización de los jugadores en pantalla, a su identificación y posterior ubicación en coordenadas reales del campo.

Según se realizaba la parte de código del proyecto, se ha ido cambiando de tecnologías disponibles las cuales en el estudio del marco teórico se pensaba que iban a ser viables. Dos de estas son los repositorios sn-calibration y tvcalib para la calibración de la cámara que en principio permitían realizar la parte de la localización del jugador en el campo real, pero que finalmente por un tema de versionado y difícil acceso a los modelos preentrenados se decidieron rechazar y optar por PnLCalib, más robusta y sencilla. También ocurrió en la parte de detección de número de dorsal, en la que se involucraba un pipeline preentrenado de detección de pose de los jugadores para interpretar más en detalle la posición del número de dorsal, cosa que resultaba demasiado compleja para el problema en cuestión y decidió rechazarse. Acabó optándose por

la solución final aplicada en el trabajo para la obtención del número del dorsal del jugador, que es de elaboración propia y surge de la necesidad de aislar el número del jugador del resto de colores, formas y ruido que hay en el marco que encuadra a cada jugador.

En cuanto a la homografía de cada frame, cabe destacar que el número de puntos que se decidió usar para inferir la posición relativa de los jugadores en el campo también surgió de la propia necesidad de encontrar puntos adicionales a los típicos (como son esquinas o puntos medios de áreas) para poder obtener al menos cuatro puntos en cada frame. Por ello, se añadieron las intersecciones de las líneas de banda con la proyección del lado largo del área grande para ambos lados, así como la intersección de las líneas de banda con la línea paralela a la de medio campo que pasa por la cúspide lateral del círculo de medio campo para ambos lados del campo. Debido a esto, también se descartó la idea de incluir una forma de calcular la posición de los jugadores cuando solo se tienen tres puntos conocida como transformación afín, ya que tampoco daba resultados fiables.

Capítulo 6. CONCLUSIONES

6.1 Conclusiones del trabajo

El resultado del trabajo cumple con el objetivo principal del proyecto ya que durante todo el desarrollo de este se aplican modelos y tecnologías de machine learning e inteligencia artificial para el análisis en profundidad de secuencias de video de partidos de fútbol. Se ha desarrollado un producto que permite obtener análisis avanzados a partir de estas secuencias y de esa manera recoger una visión profunda del juego y aportar una visión clave para la optimización del rendimiento y la estrategia de los equipos de fútbol. Este trabajo muestra que este tipo de tecnologías están disponibles para su uso de forma gratuita, lo que refuerza la idea de que cubre la brecha de obtención de análisis de fútbol significativos entre equipos grandes y pequeños.

6.2 Conclusiones personales

Durante el desarrollo del trabajo me he tenido que enfrentar mayormente en solitario a la ciencia de la visión por computador, rama de la inteligencia artificial que me atrae en especial ya que es la forma principal en la que los ordenadores son capaces de obtener información de lo que les rodea. Siempre me ha parecido interesante la futura y cada día más presente aplicación de la visión por computador a la robótica (como en los casos de robots humanoides), la conducción autónoma de coches y la vigilancia de espacios públicos y privados. Su aplicación al fútbol me ha parecido realmente interesante y prometedora.

El trabajo me ha dado una perspectiva de las principales tecnologías de uso más extendido como son YOLO o opencv y también de cómo se usan en el ámbito del fútbol, una industria que no para de crecer año a año, la cual mueve mucho dinero en el mundo y en la que cada vez más clubes fundan departamentos de análisis que basan su trabajo en los datos masivos, la inteligencia artificial y el análisis de datos.

Angel Molpeceres Mingo

La elaboración del proyecto me ha servido para desarrollar técnicas de investigación a cerca de un tema del que no había investigado nunca y me abre un mundo que desconocía como es el de la investigación.

Capítulo 7. FUTURAS LÍNEAS DE TRABAJO

En la línea del desarrollo de un producto final terminado comercializable, aparte de mejorar específicamente cada parte del pipeline de procesamiento de secuencias, se tendría que proponer una implementación en entornos distribuidos ya sea en un entorno on-sight o en la nube. Esto permitiría procesar los datos mucho más rápido, casi a tiempo real, lo que permitiría optimizar las decisiones de juego en el mismo momento en el que se está jugando.

En lo que se refiere a la depuración y mejora de las distintas partes del pipeline, también cabe hacer propuestas. Se podría desarrollar un módulo de corrección de la posición de los jugadores cuando es imprecisa utilizando su trayectoria de movimiento. Cuando se cambia entre frames o la cámara se mueve muy rápido la detección de los puntos del campo puede ser imprecisa, haciendo este fenómeno que los puntos aparezcan moviéndose lateralmente cuando los jugadores se mueven únicamente hacia delante. Inferir la trayectoria de movimiento de los jugadores podría ayudar a corregir estas pequeñas oscilaciones haciendo que el resultado final del display fuera más fluido. Ya hay estudios y papers que incluyen esta corrección, pero no ha sido añadida por cuestiones de limitación del tiempo y complejidad del presente proyecto.

También se podría trabajar más adelante en la detección del número de dorsal, ya que las predicciones son algo inconsistentes. Se podría proceder entrenando una red neuronal con un conjunto de datos extraído de las secuencias, depurar de forma distinta los recortes de los jugadores o elegir otro modelo que no fuera el OCR de Tesseract (que fue elegido por su cómoda aplicación en estos casos).

Capítulo 8. REFERENCIAS

- [1] Quality in Sport. 2025. Application and Effectiveness of Modern Technologies in Sports Training – a Literature Review. *Quality in Sport*, vol. 42, 60523. <https://apcz.umk.pl/QS/article/view/60523>.
- [2] Topend Sports. 2024. Hawk-Eye Line-Calling System – Tennis. <https://www.topendsports.com/sport/tennis/hawkeye.htm>.
- [3] St Peter's Bray. 2023. The History of Hawkeye | Pi in the Sky. <https://stpetersbray.ie/pi-in-the-sky-the-history-of-hawkeye>.
- [4] Sports Illustrated. 2015. The history of tennis umpiring: How Hawk-Eye changed the game. <https://www.si.com/tennis/2015/11/11/history-of-hawk-eye-tennis-umpiring>.
- [5] ITM Conferences. s. f. Real-Time Football Match Prediction Platform. https://www.itm-conferences.org/articles/itmconf/abs/2025/01/itmconf_dai2024_04003/itmconf_dai2024_04003.html
- [6] MECS Press. 2019. Football Match Prediction with Tree Based Model Classification. *International Journal of Intelligent Systems and Applications*, 11(7). <https://www.mecs-press.org/ijisa/ijisa-v11-n7/IJISA-V11-N7-3.pdf>.
- [7] THESAI. 2021. Developing a Reliable Hybrid Machine Learning Model for Objective Soccer Player Valuation. https://thesai.org/Downloads/Volume15No6/Paper_48-Developing_a_Reliable_Hybrid_Machine_Learning_Model.pdf
- [8] TechCentral. 2023. Sevilla FC transforms player recruitment with Watsonx generative AI. <https://www.techcentral.ie/sevilla-fc-transforms-player-recruitment-with-watsonx-generative-ai/>
- [9] Meta. 2023. How Sevilla Football Club Uses Llama and IBM Watsonx to Scout Players. <https://about.fb.com/news/2025/02/sevilla-football-club-llama-ibm-watsonx-scout-players/>
- [10] Llama. 2023. Transforming athletic scouting with AI. <https://www.llama.com/static-resource/llama-case-study-sevilla-fc/>

- [11] EAI Publications. 2024. Real Time Monitoring Research on Rehabilitation Effect of Artificial Intelligence Wearable Equipment on Track and Field Athletes. <https://publications.eai.eu/index.php/phat/article/view/5150>
- [12] Americas Publishing Group. 2024. Intelligent Integration of Wearable Sensors and Artificial Intelligence for Real-time Athletic Performance Enhancement. <https://www.americaspg.com/article/pdf/2920>
- [13] Frontiers in Sports and Active Living. 2021. Wearable Technology and Analytics as a Complementary Toolkit to Optimize Workload and to Reduce Injury Burden. <https://www.frontiersin.org/journals/sports-and-active-living/articles/10.3389/fspor.2020.630576/full>.
- [14] Folio3. 2025. Key Types of Wearable Sports Technology in 2025. <https://www.folio3.ai/blog/wearable-sports-technology/>
- [15] IEEE. 2024. MV-Soccer: Motion-Vector Augmented Instance Segmentation for CVPRW 2024.
- [16] Papers With Code. 2018. CSPDarknet53 Explained. <https://paperswithcode.com/method/cspdarknet53>.
- [17] Toshev, V.; Zlatarov, R. 2014. Multi-Targets Tracking Based on Bipartite Graph Matching. Cybernetics and Information Technologies, 14. <https://sciendo.com/article/10.2478/cait-2014-0045>
- [18] Reintech. 2022. Optical Flow for Movement Analysis in Sports with OpenCV. <https://reintech.io/blog/optical-flow-movement-analysis-sports-opencv>
- [19] Catapult. 2024. Tactical Analysis in Football: Understanding & Deploying. <https://www.catapult.com/blog/tactical-analysis-in-football>
- [20] International Journal of Computer Science in Sport. 2024. Prediction of defensive success in elite soccer using machine learning – Tactical analysis of defensive play using tracking data and explainable AI. <https://www.tandfonline.com/doi/abs/10.1080/24733938.2023.2239766>
- [21] Frontiers in Sports and Active Living. 2025. Mapping football tactical behavior and collective dynamics with artificial intelligence: a systematic review.

<https://www.frontiersin.org/journals/sports-and-active-living/articles/10.3389/fspor.2025.1569155/full>

[22] AZsportech. s. f. Analisis y tecnologia aplicados al deporte. <https://www.azsportech.com/>

[23] Fleureau, A., et al. s. f. Validity of an ultra-wideband local positioning system to assess movements in handball. <https://pubmed.ncbi.nlm.nih.gov/33343068/>

[24] Somers, V., et al. 2024. SoccerNet Game State Reconstruction: End-to-End Athlete Tracking and Identification on a Minimap. IEEE/CVF Computer Vision Workshop on Computer Vision in Sports, 2024.

https://openaccess.thecvf.com/content/CVPR2024W/CVsports/papers/Somers_SoccerNet_Game_State_Reconstruction_End-to-End_Athlete_Tracking_and_Identification_on_CVPRW_2024_paper.pdf

[25] Redmon, J., et al. 2016. You Only Look Once: Unified, Real-Time Object Detection. CVPR 2016. <https://arxiv.org/abs/1506.02640>

[26] Rodríguez, G. M. S., et al. 2024. Generalized Jersey Number Recognition Using Multi-task Learning With Orientation-guided Weight Refinement. arXiv:2406.01033. <https://web3.arxiv.org/abs/2406.01033>

[27] Gutiérrez-Pérez, M., et al. 2024. PnLCalib: Sports Field Registration via Points and Lines Optimization. arXiv:2404.08401. <https://arxiv.org/html/2404.08401v4>

[28] Veličković, P., et al. 2023. TacticAI: an AI assistant for football tactics. Nature Communications, 14:5811, 16 Oct 2023. <https://arxiv.org/abs/2310.10553>

[29] Fernández, J., Bornn, L., Cervone, D. 2021. A framework for the fine-grained evaluation of the instantaneous expected value of soccer possessions. PLOS Computational Biology. <https://link.springer.com/article/10.1101/21-05989-6>

[30] The Last Man Analytics. s. f. Voronoi Diagram. [https://thelastmanalytics.home.blog/tag/voronoi-diagram/](https://thelastmananalytics.home.blog/tag/voronoi-diagram/)

[31] Kurach, K., Andrychowicz, M., Sutskever, I. 2019. Google Research Football: A Novel Reinforcement Learning Environment. arXiv:1907.11180. <https://arxiv.org/abs/1907.11180>

[32] Cabado, B., et al. 2024/2025. Beyond the Premier: Assessing Action Spotting Transfer Capability Across Diverse Domains. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

https://openaccess.thecvf.com/content/CVPR2024W/CVsports/papers/Cabado_Beyond_the_Premier_Assessing_Action_Spotting_Transfer_Capability_Across_Diverse_CVPRW_2024_paper.pdf

[33] Owen, A., Marshall, D., Sidorov, K., et al. s. f. Assistive Sports Video Annotation: Modelling and Detecting Complex Events. Cardiff University.

<https://users.cs.cf.ac.uk/K.Sidorov/publication/sports/sports.pdf>

[34] Toolify.ai. s. f. Master the Art of Football Analysis with Self-supervised Tracking Data.

<https://www.toolify.ai/ai-news/master-the-art-of-football-analysis-with-selfsupervised-tracking-data-584103>

[35] Ruan, L., et al. 2024/2025. Accommodating Audio Modality in CLIP for Multimodal Processing. AAAI. <https://www.toolify.ai/ai-news/master-the-art-of-football-analysis-with-selfsupervised-tracking-data-584103>

[36] s. f. When to use Pruning, Quantization, Distillation and others? AI Stack Exchange.

<https://ai.stackexchange.com/questions/43054/when-to-use-pruning-quantization-distillation-and-others-when-optimizing-speed>

[37] Global Industry Analysts. s. f. Sports Analytics Market Overview. MarketResearch.com.

<https://www.marketresearch.com/Global-Industry-Analysts-v1039/Sports-Analytics-41288247/>

[38] Hidayatullah, P., et al. 2024/2025. YOLOv8 to YOLO11: A Comprehensive Architecture In-depth Comparative Review. arXiv. <https://arxiv.org/abs/2501.13400>

[39] Li, X., et al. 2020. Generalized Focal Loss: Learning Qualified and Distributed Bounding Boxes for Dense Object Detection. NeurIPS 2020. <https://arxiv.org/abs/2006.04388>

[40] Luo, X., et al. 2024/2025. Unified-IoU: For High-Quality Object Detection. arXiv.

<https://arxiv.org/html/2408.06636>

[41] Ultralytics. s. f. KalmanFilterXYAH Class Reference.

https://docs.ultralytics.com/reference/trackers/utils/kalman_filter/

- [42] Smith, R. 2007. An Overview of the Tesseract OCR Engine. ICDAR 2007. <https://ieeexplore.ieee.org/document/4376991>
- [43] Sun, K., et al. 2019. Deep High-Resolution Representation Learning for Human Pose Estimation. CVPR 2019. <https://arxiv.org/abs/1902.09212>
- [44] OpenCV. s. f. Basic concepts of the homography explained with code. OpenCV Docs. https://docs.opencv.org/4.x/d9/dab/tutorial_homography.html
- [45] MIT. s. f. Foundations of Computer Vision: Homographies. <https://visionbook.mit.edu/homography.html>
- [46] Martens, F., Dick, U., Brefeld, U. 2021. Space and Control in Soccer. Frontiers in Sports and Active Living. <https://www.frontiersin.org/journals/sports-and-active-living/articles/10.3389/fspor.2021.676179/full>
- [47] Fernandez, J., Bornn, L. 2018. Wide Open Spaces: A statistical technique for measuring space creation in professional soccer. MIT Sloan Sports Analytics Conference. https://www.researchgate.net/publication/324942294_Wide_Open_Spaces_A_statistical_technique_for_measuring_space_creation_in_professional_soccer
- [48] Spearman, L. 2018. Beyond Expected Goals. Liverpool FC. https://www.researchgate.net/publication/327139841_Beyond_Expected_Goals
- [49] Diagrama de Voronoi. <https://www.versvs.net/diagramas-de-voronoi/>
- [50] Mapa de control de campo. <https://tonyelhabr.rbind.io/posts/soccer-pitch-control-r/>
- [51] Diagrama de campo de fútbol. [https://pixabay.com/es/vectors/campo-de-fútbol-diagrama-verde-307046/](https://pixabay.com/es/vectors/campo-de-f%C3%BAtbol-diagrama-verde-307046/)

Capítulo 9. ANEXO

Todo el código elaborado puede encontrarse en el siguiente repositorio de GitHub:
<https://github.com/AngelMolpeceresMingo/TFM-Angel-Molpeceres-Mingo>.

[PÁGINA INTENCIONADAMENTE EN BLANCO]