

Estimación de la Distancia de Combate entre Boxeadores Amateurs y Posicionamiento en el Cuadrilátero mediante Visión por Computadora

Estimating Fighting Distance and Ring Positioning Between Amateur Boxers Using Computer Vision

Jose Carlos Montes González, Everardo Santiago Ramírez, René Noriega Armendáriz, Josué Domínguez Guerrero y Saúl González Campos

Instituto de Ingeniería y Tecnología, Universidad Autónoma de Ciudad Juárez, Avenida del Charro No.450 norte, Col. Partido Romero, C.P.: 32320, Ciudad Juárez, Chihuahua, México al177142@alumnos.uacj.mx, everardo.santiago@uacj.mx, rnoriega@uacj.mx, josue.dominguez@uacj.mx, saugonza@uacj.mx

Fecha de recepción: 18 de septiembre de 2024

Fecha de aceptación: 27 de marzo de 2025

Resumen. La visión por computadora puede ser empleada para analizar objetivamente el rendimiento de los boxeadores amateurs. A diferencia de los métodos tradicionales que se basan en observaciones subjetivas, este trabajo presenta una aplicación para estimar de manera más precisa la distancia de combate y el posicionamiento de los boxeadores dentro del cuadrilátero. Utilizando un modelo de aprendizaje profundo pre-entrenado, el sistema procesa videos de peleas y calcula la distancia entre los combatientes y su ubicación dentro del cuadrilátero. Los resultados obtenidos demuestran la precisión y fiabilidad de este enfoque, lo que lo convierte en una herramienta útil para entrenadores y analistas deportivos. Esta aplicación puede ayudar a identificar fortalezas y debilidades, mejorar el entrenamiento y mejorar el rendimiento en futuras competencias.

Palabras clave: Visión por computadora, Aprendizaje profundo, Boxeo.

Summary. Computer vision can be used to objectively analyze the performance of amateur boxers. Unlike traditional methods that rely on subjective observations, this work presents an application to more accurately estimate the fight distance and positioning of boxers within the ring. Using a pre-trained deep learning model, the system processes fight videos and calculates the distance between the combatants and their location within the ring. The results obtained demonstrate the accuracy and reliability of this approach, making it a useful tool for coaches and sports analysts. This application can help identify strengths and weaknesses, improve training, and enhance performance in future competitions.

Keywords: Computer Vision, Deep Learning, Boxing.

1 Introducción

La visión por computadora se ha convertido en una herramienta indispensable en el análisis de videos deportivos debido a su capacidad para proporcionar momentos detallados y objetivos que son difíciles de obtener mediante observación convencional. En el ámbito del boxeo amateur, la capacidad para analizar el posicionamiento en el cuadrilátero y la distancia de combate son factores críticos que pueden influir significativamente en el desempeño de los boxeadores. El posicionamiento en el cuadrilátero se refiere a la ubicación estratégica y al movimiento de un boxeador dentro del mismo durante un combate, siendo este fundamental en su capacidad para atacar, defenderse y controlar el ritmo del combate.

Como se puede observar en la Fig. 1, el cuadrilátero se divide en tres zonas (Z1, Z2 y Z3). Z1 corresponde al centro, Z2 a las áreas adyacentes a los lados (o esquinas) y Z3 a las regiones opuestas a los lados.

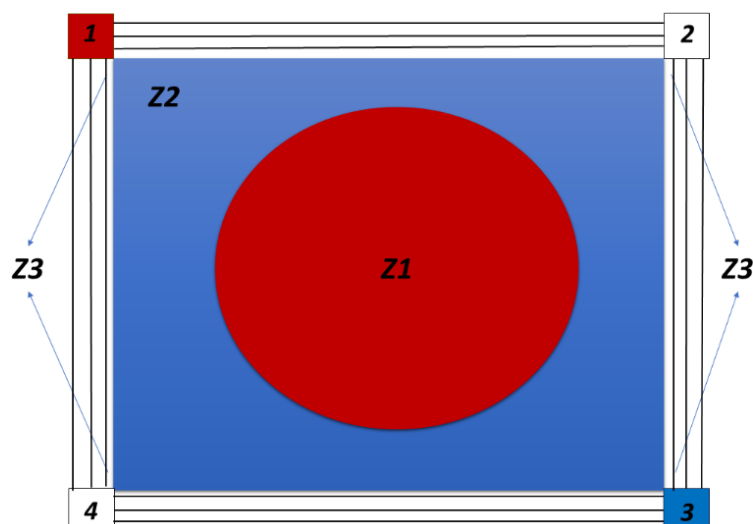


Figura 1. Zonas que conforman un cuadrilátero de boxeo.

La distancia en el boxeo es un factor crucial que determina el éxito de un combate. Las tres distancias principales —larga, media y corta— ofrecen a los boxeadores distintas opciones tácticas y técnicas. A mayor distancia, los púgiles pueden emplear golpes rápidos y directos, así como ganchos al hígado y *uppercuts*. En la distancia media, se amplía el repertorio de golpes, permitiendo el uso del cruzado. La distancia corta, por su parte, favorece los ataques cuerpo a cuerpo, donde la precisión y la potencia son fundamentales para conectar los golpes.

Ante la necesidad de contar con herramientas tecnológicas más precisas en el entrenamiento del boxeo, hemos fusionado el conocimiento empírico de los entrenadores con los avances en visión por computadora. El resultado es un prototipo que analiza videos de combates amateurs, proporcionando datos cuantitativos sobre variables clave como la distancia de combate y el posicionamiento en el ring. Al visualizar estos datos en tiempo real, los entrenadores pueden complementar su experiencia y optimizar el proceso de entrenamiento.

2 Estado del arte

El deporte es un componente fundamental de la cultura humana que abarca una amplia gama de actividades físicas que van desde el ejercicio recreativo hasta la competición profesional. El análisis en el deporte ha experimentado una transformación revolucionaria en las últimas décadas debido a los avances tecnológicos y la creciente disponibilidad de datos. Lo que solía ser una evaluación subjetiva y basada en observaciones ha evolucionado hacia un enfoque altamente objetivo, recopilando y estudiando conjuntos de datos provenientes de múltiples fuentes. La información recopilada se procesa utilizando herramientas tecnológicas avanzadas, las cuales permiten identificar patrones, tendencias y elementos clave dentro del juego.

En cuanto a la investigación de la visión artificial aplicada al deporte, es importante abordar la falta de datos e imágenes específicas para ciertos deportes, como por ejemplo el ráquetbol. En [1] se menciona una recopilación de imágenes del ráquetbol, encontrando poco volumen de información. Como solución se utiliza una cámara de alta resolución para capturar imágenes de la cancha y se utilizan técnicas de inteligencia artificial, como la detección de objetos, en la que el algoritmo detecta a los jugadores en las imágenes. Se detecta el movimiento de los jugadores y se utiliza un algoritmo de clasificación de objetos identificando así las técnicas de los jugadores. Se demuestra que el reconocimiento de acciones humanas en el deporte es una tarea difícil de llevar, pero con bastante potencial de mejora.

Se expuso en [2] que el problema fundamental en el deporte es el poco uso de inteligencia artificial, ya que esta aún no ha madurado lo suficiente. Debido a esto se propone el uso de una cámara de alta velocidad para capturar imágenes a 60 fotogramas por segundo permitiendo así capturar las imágenes de los jugadores con mayor detalle sobre su movimiento donde se obtiene la variabilidad de las acciones, también se propuso el uso de un algoritmo de detección y seguimiento de objetos para la identificación y rastreo de jugadores en videos. Se recomendaron métodos que han demostrado ser efectivos en el reconocimiento de acciones humanas en deportes.

Se examinó en [3] la problemática sobre los videos deportivos que son largos, lo que los vuelve muy tediosos de ver en su totalidad. Los espectadores solicitan un resumen de los momentos más importantes de una forma más clara y eficiente, sin dejar de lado el audio y video al procesarlo. Se creó un método para generar resúmenes de

videos deportivos. Los resultados arrojados muestran que es posible dar un resumen de los videos deportivos siendo fiel a su contenido original, además, los espectadores de un grupo de prueba concluyeron que los resúmenes eran más fáciles de ver y comprender que los videos originales.

Con respecto a la falta de visualización aceptable y de los análisis deportivos de manera efectiva, en [4] se menciona que los datos se utilizan en diferentes ámbitos. Cada una tiene su forma específica de visualización, por tanto, la solución propuesta fue una taxonomía de técnicas de visualización. Dicha taxonomía es fundamental para el entendimiento y los usuarios interesados en descubrir más del tema y términos como: tipo de datos, tareas de visualización, técnicas y sus respectivas definiciones. Además, estos se clasifican dependiendo de su área, estadísticas, movimiento, video y gráficos. La propuesta contribuye a una comprensión y aprovechamiento de la información en el ámbito deportivo.

Para los videos deportivos, donde se interponen los movimientos de la cámara, las oclusiones y el desenfoque, en [5] se menciona que se requiere de un algoritmo que pueda resolver estas necesidades y que sea adecuado para trabajarlo en tiempo real. Por lo tanto, se propuso un algoritmo de seguimiento para aplicaciones deportivas, el cual usa el método de Kalman y se ha optimizado en dispositivos móviles para su portabilidad. Este algoritmo destaca debido a la predicción de la posición y el movimiento del objeto de cada fotograma, utilizando la imagen actual para actualizar las imágenes. El algoritmo se ejecuta en dispositivos móviles, el cual siguió de manera precisa y efectiva objetos en videos deportivos, esto con un 95% de éxito y un tiempo de 10 milisegundos por fotograma.

En [6] se presentó lo difícil que es distinguir los objetos de los atletas en videos con movimiento a través de una cámara, esto debido a que los deportistas suelen ser similares en su forma física y es difícil detectar algunos objetos que quedan bloqueados. El movimiento de la cámara puede ser de manera impredecible, lo que puede tener detecciones erróneas y asociar detecciones de objetos. La solución empleada fue combinar dos algoritmos: Deepsort y Yolo V4. El primero se utilizó para crear un algoritmo que trabaja la re-identificación de objetos y utiliza aprendizaje profundo para asociar las detecciones de los objetos localizados, mientras que Yolo V4 es un algoritmo de detección de objetos que trabaja de manera rápida y precisa. Este método es robusto al manejar el seguimiento de los objetos, ya sea que tengan apariencia similar o escaneando la detección parcial del objeto y que es fácil de implementar.

3 Metodología

En el diagrama de la Fig. 2 se ilustra gráficamente el proceso implementado en la aplicación que se reporta en este documento para analizar los combates de boxeo. La entrada es un video en donde se vean de manera vertical ambos boxeadores, tal y como se muestra en la Fig. 3. En este video tanto la cámara como el cuadrilátero permanecen fijos. Una vez recibido el video de entrada, se obtiene la primera imagen en donde se asume que aparece un cuadrilátero. El usuario marca las cuatro esquinas del cuadrilátero, y también traza una circunferencia para delimitar el área de la Z1. Estas marcas son usadas en cada fotograma para identificar automáticamente la zona en que se encuentran los boxeadores.

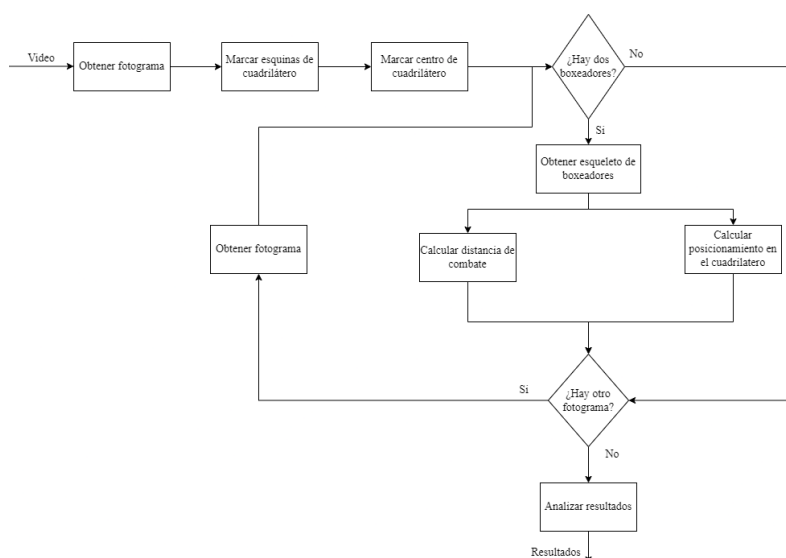


Figura 2. Diagrama de bloques para el análisis de videos de combates boxísticos.

Si el fotograma actual contiene dos boxeadores, entonces se obtiene el esqueleto de cada uno de ellos. El esqueleto humano se obtiene usando la biblioteca de visión por computadora Detectron2. Para esto se detectan puntos semánticos vitales, que en caso del cuerpo humano se refieren, principalmente, a las articulaciones. Ejemplo de articulaciones son hombros, codos, muñecas, rodillas, tobillos, entre otros. Estos puntos clave se asocian para formar el esqueleto. En la Fig. 3 se puede observar ejemplos de esqueletos de boxeadores.

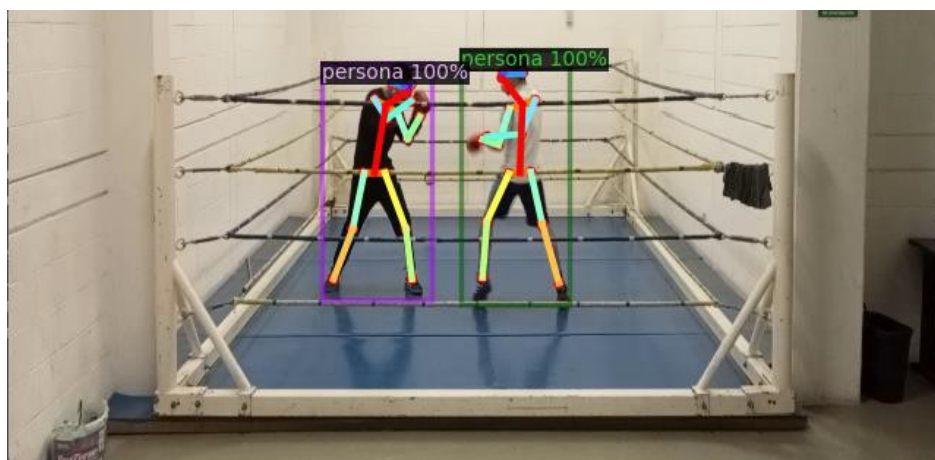


Figura 3. Esqueletos de los boxeadores en un combate.

Una vez obtenidos los esqueletos de los cuerpos de los boxeadores, se calcula la distancia entre los combatientes. Para esto, de todos los puntos clave, de cada boxeador se seleccionan solo los que corresponden a las rodillas, tobillos, muñecas y cabeza. Luego, se calculan las distancias euclidianas dr , dt , dm y dc entre las rodillas, tobillos, muñecas y cabeza de los boxeadores, respectivamente. Después, se calcula el promedio de las distancias $D = dr+dt+dm+dc$ y se determina el tipo de distancia:

$$\begin{aligned}
 & \text{corta si } D \leq \overline{u_1} \\
 \text{Tipo de distancia} &= \{ \text{media si } u_1 < D \leq u_2 \\
 & \text{larga si } D > \overline{u_2}
 \end{aligned}$$

donde u_1 y u_2 , son valores umbrales que delimitan las zonas Z_1 y Z_2 , respectivamente. Para definir el posicionamiento en el cuadrilátero de cada boxeador, se usan las ubicaciones del tobillo izquierdo ti y tobillo derecho td . Si estos se encuentran dentro de la circunferencia marcada por el usuario, entonces el boxeador se ubica en Z_1 . No obstante, si los tobillos se ubican fuera de esa circunferencia, pero dentro del cuadrilátero, entonces el boxeador se ubica en la Z_2 . En otro caso, el boxeador se encontraría en la Z_3 , lo cual indicaría que se encuentra fuera del cuadrilátero. Es fácil notar que para la distancia de combate se requiere información de ambos boxeadores, mientras que el posicionamiento en el cuadrilátero puede ser calculado de manera individual.

4 Resultados experimentales

La aplicación que implementa lo descrito en la sección anterior fue programada en el lenguaje de programación Python debido a su versatilidad y claridad, además de que es ideal para la ciencia de datos y aprendizaje máquina. Para el procesamiento de imágenes, se empleó la biblioteca de visión artificial OpenCV. Esta biblioteca se usó principalmente para la lectura y escritura de imágenes digitales.

Como se mencionó anteriormente, un elemento importante para el desarrollo del prototipo propuesto fue la biblioteca Detectron2 que implementa algoritmos de segmentación y detección. Esta biblioteca se usó para la detección de articulaciones del cuerpo de los boxeadores y formar así su esqueleto. No se requirieron datos nuevos para Detectron2 ya que esta viene con pesos previamente entrenados. La aplicación fue desarrollada usando las herramientas anteriores en el entorno Anaconda y en el sistema operativo Windows 11.

En la Figura 4 se muestra el resultado de la ejecución del prototipo propuesto en el primer fotograma. Las esquinas del cuadrilátero seleccionadas se muestran en círculos color verde. La circunferencia en rojo indica la Z_1 , la cual tiene que estar en el centro del cuadrilátero. De este círculo se deducen las demás zonas por donde se posicionan los boxeadores. Como se puede notar, la señalización precisa de estos puntos es importante para la correcta determinación del posicionamiento de los peleadores.

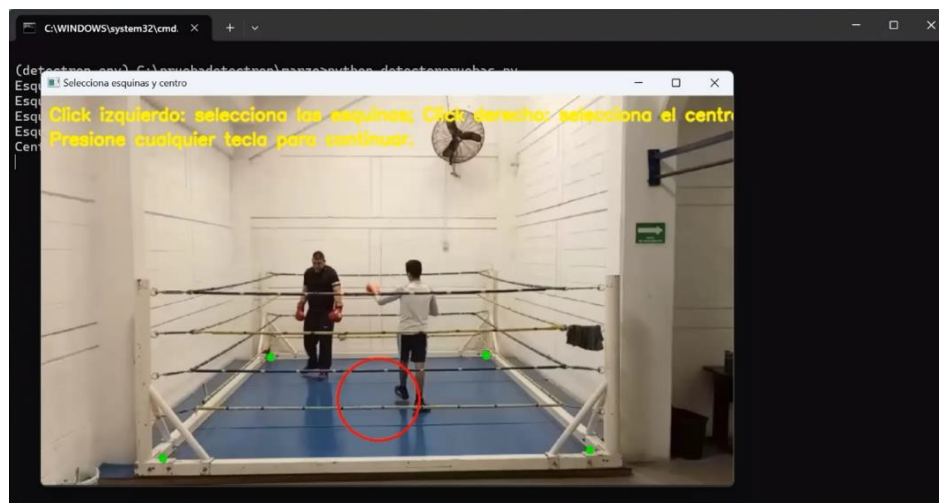


Figura 4. Selección de esquinas y zona central del cuadrilátero.

Una vez marcadas las esquinas y el centro del cuadrilátero, el prototipo empieza a obtener en cada fotograma el esqueleto de los boxeadores. Con los esqueletos se calcula la distancia de combate, que como se puede observar en la Fig. 5, conforme se avanza en el análisis del video, se muestra al usuario el porcentaje de tiempo que lleva el boxeador de interés en alguna zona del cuadrilátero y el número de fotogramas que ha realizado alguna distancia de combate. Esta forma de presentar el resultado permite al usuario ver los resultados en tiempo de reproducción. Para conocer mejor el funcionamiento del prototipo, se hicieron dos experimentos.

Este análisis permitió obtener información de un total de 196 fotogramas, de los cuales, los cuadros en donde el boxeador peleaba a distancia corta fueron 10, en distancia media 18 y distancia larga fueron 168. Para una mejor comprensión del funcionamiento del prototipo, se tomó solo un fragmento de 24 fotogramas. De estos, en solo 14 cuadros se detectaron a los dos boxeadores de manera correcta y se pudieron calcular la distancia de combate y el posicionamiento en el cuadrilátero. Esto sugiere que no todos los fotogramas de los videos contienen información útil, sin embargo, la mayoría de éstas contienen una cantidad sustancial de datos utilizables para extraer información significativa sobre la pelea de boxeo.

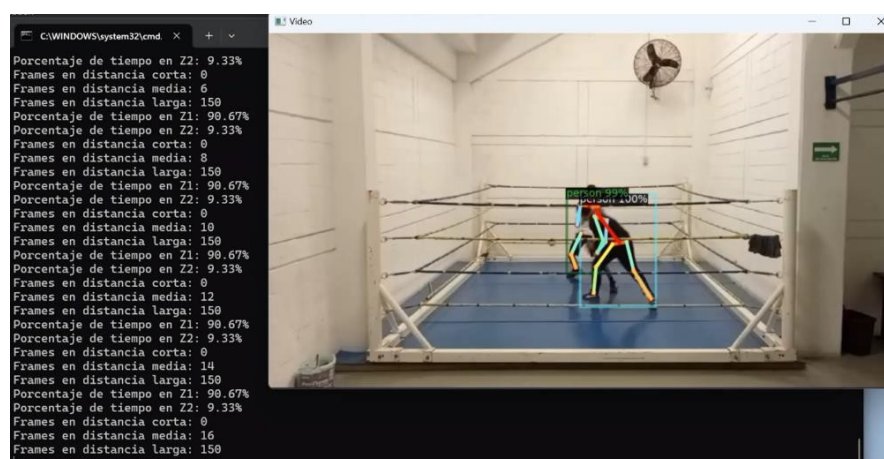


Figura 5. Resultados del prototipo propuesto durante la ejecución del prototipo.

El primer experimento tuvo como objetivo evaluar la capacidad del prototipo para calcular la distancia de combate entre dos boxeadores en un video de un combate real. El fragmento utilizado tuvo una duración 35 segundos donde se sintetizó un combate de 3 minutos con 16 segundos. En dicho video se puede observar a los boxeadores en movimiento, donde uno de ellos se encuentra de espaldas y se encuentran lanzando golpes en un cuadrilátero. Para analizar la distancia de combate, se utilizó un modelo de visión por computadora para seguir los puntos clave de los boxeadores en cada fotograma del video. A partir de esta información se calculó la distancia entre los boxeadores en cada momento del combate.

En el segundo experimento, se dio prioridad al análisis de la distribución del espacio de los boxeadores y su control del territorio dentro del cuadrilátero. Para ello, se usó un fragmento de video de un combate real de una duración de 46 segundos. Los boxeadores se mantuvieron en movimiento y lanzaron golpes en las zonas del

cuadrilátero. El prototipo se enfocó en seguir los tobillos para calcular la posición de ambos boxeadores en el cuadrilátero. Tomando solo en cuenta el fragmento de 18 cuadros, de los cuales se obtuvieron 11 que detectaron a los dos boxeadores de manera correcta. Los fotogramas efectivos fueron 11, los cuales permitieron obtener información relevante sobre la estrategia espacial de los boxeadores y su capacidad para controlar el territorio.

Ambos experimentos permiten observar la capacidad del prototipo para analizar el comportamiento de los boxeadores en un combate amateur. En el experimento 1, se evidencia la capacidad del prototipo para realizar mediciones de distancia de combate entre dos boxeadores en un cuadrilátero de boxeo. Por su parte, el experimento 2 se centró en el análisis del posicionamiento de los boxeadores en el cuadrilátero, dando resultados que permitieron comprender la distribución espacial de los boxeadores y su control del territorio durante el mismo. En conjunto, estos experimentos validan la efectividad del prototipo como herramienta para el análisis del movimiento en el boxeo, brindando nuevas posibilidades para su aplicación en el entrenamiento, el análisis de competencias y la investigación científica en el ámbito del boxeo en un futuro. Sin embargo, los resultados son precisos solo cuando un boxeador está a la izquierda y el otro a la derecha. Cuando se mueven hacia sus lados, los resultados pierden su precisión gradualmente. Una solución pudiera ser el empleo de la información de profundidad.

5 Conclusiones y trabajo a futuro

En este estudio, se ha presentado un prototipo innovador diseñado para calcular de manera precisa la distancia de combate y el posicionamiento de los boxeadores dentro del cuadrilátero. Los resultados obtenidos demuestran la viabilidad de analizar combates de boxeo empleando técnicas avanzadas de visión artificial. Esta información objetiva y detallada puede ser una herramienta invaluable para entrenadores, quienes podrán tomar decisiones estratégicas y personalizadas en el entrenamiento de sus atletas. El prototipo propuesto tiene el potencial de revolucionar el entrenamiento de boxeadores amateur al proporcionar datos cuantitativos sobre el desempeño de cada deportista. Estos datos permitirán a los entrenadores identificar con mayor precisión las fortalezas y debilidades de sus boxeadores, diseñar planes de entrenamiento altamente personalizados y adaptados a las necesidades individuales de cada atleta y desarrollar estrategias de combate efectivas y crear contra estrategias para enfrentar a diversos oponentes.

Como trabajo futuro, se plantea la necesidad de emplear información de profundidad para el cálculo de los datos y realizar pruebas más exhaustivas para validar la precisión de los cálculos del prototipo en una variedad de escenarios. Además, se buscará desarrollar una interfaz gráfica de usuario intuitiva que facilite la interpretación de los resultados por parte de los entrenadores. Finalmente, se explorará la posibilidad de procesar múltiples videos en paralelo para aumentar la eficiencia del sistema.

Referencias

- [1] M. Baclig, N. Ergezinger, Q. Mei, M. Gül, S. Adeeb, and L. Westover, “A deep learning and computer vision based multi-player tracker for squash,” *Applied Sciences*, vol. 10, p. 8793, 12 2020.
- [2] K. Host and M. Ivašić-Kos, “An overview of human action recognition in sports based on computer vision,” *Heliyon*, vol. 8, p. e09633, 06 2022.
- [3] Khan, J. Shao, W. Ali, and S. Tumrani, “Content-aware summarization of broadcast sports videos: An audio–visual feature extraction approach,” *Neural Processing Letters*, vol. 52, 12 2020.
- [4] M. Du and X. Yuan, “A survey of competitive sports data visualization and visual analysis,” *Journal of Visualization*, vol. 24, 08 2020.
- [5] V. Vovk, S. Skuratovskyi, P. Vyplavin, and I. Gorovyi, “Light-weight tracker for sports applications,” in *2019 Signal Processing Symposium (SPSymposium)*, pp. 251–255, 2019.
- [6] Y. Zhang, Z. Chen, and B. Wei, “A sport athlete object tracking based on deep sort and yolo v4 in case of camera movement,” in *2020 IEEE 6th International Conference on Computer and Communications (ICCC)*, pp. 1312–1316, 2020.