

Sistema de inteligencia artificial para la asistencia y corrección de técnicas de deportistas de alto rendimiento por visión artificial

Artificial intelligence system for assisting and enhancing high-performance athletes' techniques using computer vision.

Carlos-Alfonso Trejo-Villanueva ^a, Derlis Hernández-Lara ^a, Emmanuel-Tonatiuh Juárez-Velázquez ^a, Cinthia-Estela Trejo-Villanueva ^a

^a Ingeniería informática, Tecnológico de Estudios Superiores de Ecatepec, C.P. 55210, Ecatepec de Morelos, Estado de México, México

Resumen

En la actualidad, la visión artificial (VA) está experimentando un auge significativo gracias a los avances en inteligencia artificial (IA), aprendizaje automático (machine learning), la creciente disponibilidad de datos y potencia de cómputo. Las principales áreas de desarrollo de la IA son la industria manufacturera, el área de la salud, la automoción, seguridad y vigilancia, agricultura, deporte y otras. En referencia al deporte de alto rendimiento, se establece que una de las grandes limitantes es la relación entre la cantidad de entrenadores capacitados y la cantidad de deportistas que atiende, entendiéndose que el tiempo de atención a cada deportista se ve disminuido cuando aumenta la cantidad de personas con las que se entrena, generando un retraso en sus avances deportivos y posibles lesiones debido al desarrollo de técnicas erróneas a causa de esta desatención. El presente trabajo desarrolla un sistema de asistencia para deportistas de alto rendimiento basado en inteligencia artificial, para el perfeccionamiento de técnicas y corrección de errores en su ejecución. Se basa en algoritmos de visión artificial que permite establecer técnicas de entrenamiento y mecánicas de ejecución de ciertos ejercicios, haciendo uso de video en tiempo real en el que se informa a los deportistas de la correcta o deficiente ejecución de su técnica. Los resultados y conclusiones describen los pormenores del desarrollo aplicados a técnicas de deportistas en el patinaje libre y los alcances de la implementación del prototipo.

Palabras clave: Deporte de alto rendimiento, Inteligencia Artificial, Media Pipe Pose, Python, Visión artificial.

Abstract

Currently, computer vision (CV) is experiencing significant growth due to advances in artificial intelligence (AI), machine learning, the increasing availability of data and computing power. The main areas of development of IA are the manufacturing industry, healthcare, automotive, security and surveillance, agriculture, sports, and many others. Regarding high-performance sports, it is established that the major problem is the ratio between the number of trained coaches and the number of athletes they attend to, understanding that the attention time for each athlete is diminished when the number of people being trained increases, resulting in a delay in their athletic progress and possible injuries due to the development of erroneous techniques due to such inattention. This work develops an assistance system for high-performance athletes based on artificial intelligence, for the enhancing of techniques and correction of execution errors. It is based on computer vision algorithms that allow establishing training techniques and execution mechanics of some exercises, using real-time video in which athletes are informed of the correct or deficient execution of their technique. Results and conclusions describe the details of the development applied to techniques of athletes in skating and the scope of the prototype implementation.

Keywords: High-performance sport, Artificial Intelligence, Media Pipe Pose, Python, Computer vision

1. Introducción

En el mundo de los deportes de alto rendimiento, los entrenadores y preparadores físicos son responsables del entrenamiento de los deportistas teniendo como alguna de sus

*Autor para la correspondencia: josejuan@dominio1.edu.mx

Correo electrónico: carlostrejo@tese.edu.mx (Carlos-Alfonso Trejo-Villanueva), dderlis-lara@tese.edu.mx (Derlis Hernández-Lara), emmanuel.juarez@tese.edu.mx (Emmanuel-Tonatiuh Juárez-Velázquez), cinthia.trejo@tese.edu.mx (Cinthia-Estela Trejo-Villanueva).

prioridades la de corregir las técnicas y posturas de dichos deportistas, entendiendo que la teoría moderna del entrenamiento refiere a un proceso sistemático de integración multidisciplinaria (Chunguang, 2023). No importa el deporte de alto rendimiento que se mencione, la especialización de técnicas correctas de ejecución o mecánicas de motricidad son imperantes para alcanzar los mejores resultados en ellos. Todo ejercicio provoca procesos de adaptación aguda, esenciales para ajustar las funciones corporales al nivel correspondiente de metabolismo energético, siendo necesarios ajustes en cada deportista para prevenir alteraciones dañinas en el medio interno de su organismo, que permiten mejorar el rendimiento en el ejercicio y mejorar el nivel de desempeño (Virus, 2017). Sin embargo, en muchas ocasiones, la saturación de los entrenadores respecto al número de deportistas que atienden no permite otorgar el tiempo necesario para la identificación los ajustes técnicos de errores técnicos y la personalización del entrenamiento para reflejar la capacidad y respuesta son únicas para cada atleta (Kasper 2019). Al tener demasiados deportistas a su cargo, obligan al entrenador a disminuir el tiempo dedicado a cada uno para realizar las correcciones necesarias o dar las indicaciones pertinentes. Una técnica incorrecta puede aumentar la posibilidad de lesiones, reducir el rendimiento o incluso impedir que el deportista alcance su máximo potencial (Luckstead, 2002). El uso de la Inteligencia artificial ha ayudado a muchas áreas a mejorar la eficiencia y en el caso del problema descrito previamente puede ayudar a resolver este problema al brindar una herramienta objetiva y precisa para la identificación de las técnicas adecuadas de entrenamiento.

2. Inteligencia artificial en los deportes

La inteligencia artificial (IA), es un campo que combina la ciencia informática y los conjuntos de datos robustos para permitir la resolución de problemas. El concepto de Inteligencia Artificial está alterando positivamente el deporte y elevándolo a un nivel de éxito sin precedentes (Nadikattu, 2020). Así, la IA está transformando rápidamente el mundo del deporte, desde el entrenamiento y la preparación física hasta el análisis del rendimiento y la experiencia del espectador. La IA está revolucionando el ámbito del entrenamiento deportivo al emplear técnicas avanzadas de aprendizaje automático y algoritmos computacionales (Bodemer, 2023). Algunas de las aplicaciones más destacadas de la IA en los deportes son:

- Análisis del Rendimiento
- Entrenamiento y Preparación Física
- Experiencia del Espectador
- Gestión Deportiva

El análisis biomecánico asistido por IA estudia el movimiento de los atletas con gran detalle, identificando deficiencias técnicas, riesgos de lesiones y áreas de mejora,

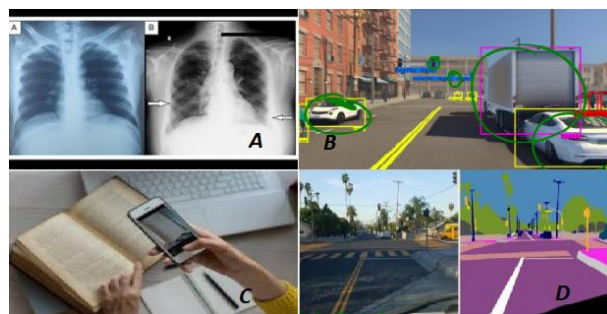
para lo cual se suelen utilizar técnicas de visión artificial (visión por computadora) para capturar y analizar la motricidad de los deportistas. Estudios recientes se centran en la utilización de técnicas de IA para extraer automáticamente poses humanas de flujos de video, con el objetivo de evaluar objetivamente los gestos de los atletas y mejorar su rendimiento deportivo (Annino, 2023). El análisis automático de vídeos en el ámbito deportivo se presenta como una solución viable para satisfacer la demanda de información variada por parte de aficionados y profesionales (Naik, 2022). Los sistemas de seguimiento basados en cámaras y sensores recopilan datos sobre la posición, la velocidad, la aceleración, la distancia recorrida y otros parámetros en tiempo real. El objetivo de estos sistemas pretende analizar datos biomecánicos y de rendimiento para identificar factores de riesgo de lesiones y recomendar medidas preventivas y así mejorar la eficiencia, eficacia, rendimiento y resultados de los deportistas en su proceso de preparación y entrenamiento.

3. Visión Artificial

La visión artificial (VA) es un campo de la inteligencia artificial que se enfoca en hacer que las computadoras interactúen con el mundo visual de forma similar a como lo hacemos los humanos y así puedan integrar y analizar información visual del mundo que las rodea, en gran medida, estos avances han sido impulsados por sistemas base potentes (He, 2017).

Esta disciplina científica, se basa en redes neuronales convolucionales (CNN), que han demostrado ser modelos eficaces para abordar una variedad de tareas visuales (Hu, 2018) que buscan conseguir que las computadoras sean capaces de comprender imágenes y actuar de cierta manera ante ellas. Las aplicaciones en las que se usa la VA son básicamente tareas de:

- Clasificación
- Detección
- Optical Character Recognition (OCR)
- Segmentación



• Figura 1: Aplicaciones de la VA. A) Clasificación, B) Detección, C) OCR, D) Segmentación

Dependiendo de la aplicación que se pretende desarrollar se utiliza una de estas técnicas, buscando identificar ciertos parámetros de imagen que permita realizar un proceso específico y automático.

4. Método

En esta investigación se realizó un análisis de imágenes recopiladas de deportistas de alto rendimiento dedicados al patinaje de velocidad. La metodología de trabajo esta descrita en la Figura 2.

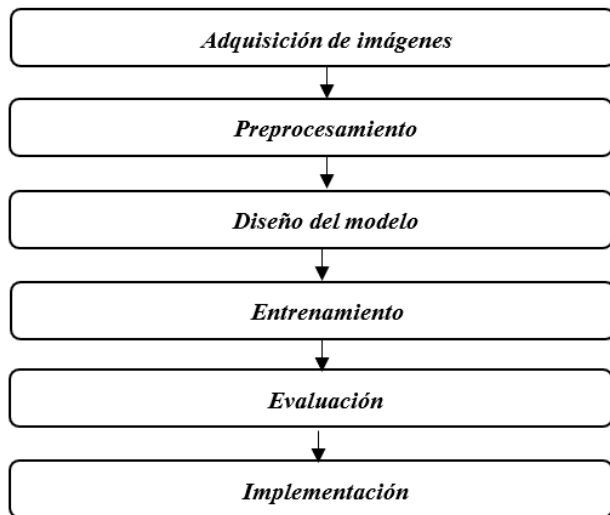


Figura 2: Metodología de trabajo empleada.

4.1. Adquisición de imágenes

El banco de imágenes utilizado para la identificación de posturas es generado por el equipo de investigación específicamente para el estudio y se compone de la captura de 6000 imágenes de dos diferentes clasificaciones basado en las posiciones en las que los patinadores se presentan en la imagen (3000 para cada clasificación). Las imágenes son validadas por los expertos en el área deportiva para ser etiquetadas correctamente en las clasificaciones establecidas que son:

1. Mala postura
2. Buena postura

Las posturas son definidas con las marcas al esqueleto que se sobrepone a la imagen del patinador, guardando referencia entre los puntos y ángulos formados por las rodillas, los tobillos, la cadera, los hombros y la cabeza. Se define una mala postura cuando los ángulos de las rodillas quedan por dentro de los puntos de la cadera o cuando los puntos de los tobillos están más allá de los puntos de las rodillas. Así mismo la referencia entre los hombros y la cadera deben estar alineados y los puntos de la cabeza no deben presentar una inclinación en la misma para ser considerada una buena postura.

Las imágenes son capturadas con una cámara GoPro Hero12 con estabilizador de captura HyperSmooth 6.0 con Bloqueo del horizonte de 360° que puede grabar video o tomar imágenes de 5.3K y añade HDR (Alto Rango Dinámico) y puede capturar fotografías con fotogramas de 24,7 MP a partir de videos.

La captura de las fotografías se configura a una resolución de 1920 x 1080 px, por ser su resolución mínima. Estas imágenes se toman en un entorno semi controlado con iluminación, ángulo de toma, persona fotografiada y escenario similares, para dar una certidumbre de clasificación pretendida. La figura 3 muestra un ejemplo de las tomas capturadas.



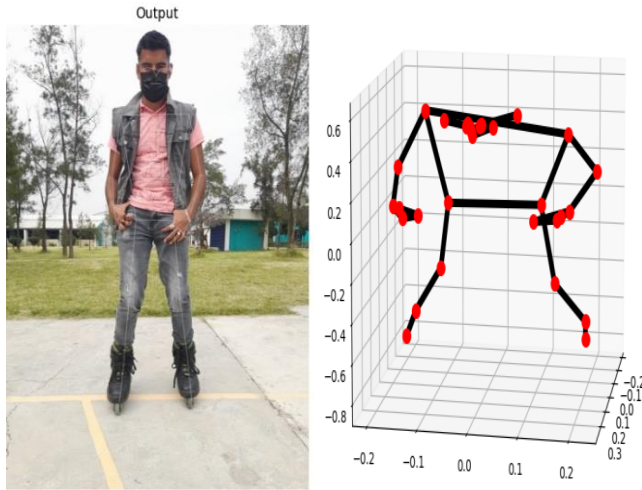
Figura 3: Ejemplo de imagen capturada.

4.2. Preprocesamiento

Se realizó diferentes tareas de procesamiento de las imágenes del banco obtenido para hacer más eficiente el procesamiento del modelo de IA.

- Redimensionamiento de imagen. Se redimensiona las imágenes cambiando la resolución original a una de 150 x 150 px, para que el entrenamiento de la red sea más rápido y eficiente. Cabe notar que las imágenes fueron revisadas para verificar que pudiera ser reconocidos las características que definirían las clasificaciones esperadas.
- Se implementa la biblioteca Media Pipe Pose (MP Pose), que es una parte de conjunto integral de soluciones de machine learning para aplicaciones multimedia, utilizada para agregar a las imágenes un esqueleto compuesto por puntos asignados a los hombros, los codos, las muñecas, la cadera, las rodillas y los tobillos de los patinadores. Estos puntos servirán para la definición de las clasificaciones establecidas en el diseño del proyecto. La figura 4 muestra la inserción del esqueleto en la imagen

- Disminución de canales de imagen. Las imágenes se transforman para que solo manejen matrices de canales RGB., ya que la coloración de las imágenes no es un parámetro relevante en la clasificación buscada. No se trabaja con imágenes en escala de grises por la interpretación del preprocesamiento anteriormente descrito
- Estandarización de los píxeles. Las 3 matrices de las imágenes definen los píxeles en valores que van de 0 a 255 con lo que las operaciones implícitas en la red pueden ser costosas computacionalmente, por lo que se aplica una estandarización para transformar los datos a valores en el rango de 0 a 1 y hacer así que las operaciones sean con valores pequeños y disminuir así el tiempo de entrenamiento y clasificación.



• Figura 4: Imagen capturada con esqueleto generado por la librería BlazePose.

4.3. Diseño del modelo y entrenamiento

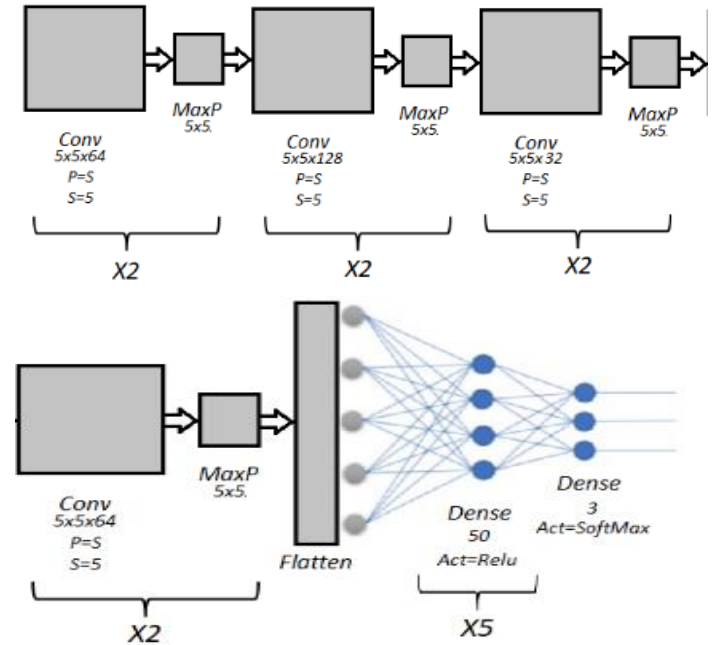
En la experimentación se prueban diferentes configuraciones de arquitectura de modelos cambiando los parámetros: número de capas, funciones de activación, índice de aprendizaje, número de épocas, tamaño de filtro y número de filtros, tamaño de batch, padding y stride. Se realizan 500 ejercicios de experimentación con el cambio de diferente meta parámetros, de estas experimentaciones se presenta en la Tabla 1 el modelo que proporcionó los mejores resultados con el mínimo de requerimientos computacionales.

Tabla 1: Meta parámetros del modelo de clasificación utilizado

Parámetro	Valor
Número de capas	24
Tipos de capas	Convolutacional, MaxPooling, Flatten y Dense
Funciones de activación	Relu, Softmax
Índice de aprendizaje	0.0003
Épocas	60

Tamaño de batch	100
Optimizador	Adam
Pérdida	Categorical crossentropy
Métrica	Accuracy

El banco de 6000 imágenes preprocesadas se divide en 70% de imágenes para el entrenamiento (training) y 30% de imágenes para las pruebas (test). El diagrama de la arquitectura que presentó mejores resultados se muestra en la Figura 5.



• Figura 5: Diagrama de la arquitectura del modelo.

El resumen de las capas del diseño del modelo de la arquitectura creada se detalla en la Tabla 2.

Tabla 2: Modelo detallado

Capa	Tipo	Parámetros
1	Convolutacional X8	Kernel (64x2, 128x2, 32x2, 64x2) Ventana (5x5) Padding (same) Stride (5)
2	MaxPooling X8	Pool size (5x5)
3	Flatten	
4	Dense X5	Activation (Relu) No. neuronas (50)
5	Dense	Activation (SoftMax) No. neuronas (3)

4.4. Evaluación e implementación

En la implementación se realizan las pruebas de evaluación en experimentos en los que se cambian parámetros como número de capas, funciones de activación, índice de aprendizaje, número de épocas, tamaño de filtro y número de filtros, tamaño de batch, padding y stride. De estos experimentos, se presentan en este trabajo los valores de la arquitectura que mostró los mejores resultados considerando las magnitudes del error (Categorical Cross Entropy) y de la exactitud (Accuracy). Los valores de pérdida y precisión se muestran en las gráficas de la siguiente sección con el progreso del entrenamiento en el transcurso de las épocas.

5. Resultados

Para evaluar los resultados del modelo diseñado se utilizó la métrica Accuracy y la función de pérdida Categorical Cross Entropy, con lo que se alcanzó los resultados mostrados en la Tabla 3.

Tabla 3: Valores de pérdida y precisión alcanzados

Etap	Parámetro	Valor
<i>Entrenamiento</i>	<i>Accuracy</i>	<i>0.8297</i>
	<i>Loss</i>	<i>0.4751</i>
<i>Validación</i>	<i>Val- Accuracy</i>	<i>0.8187</i>
	<i>Val - Loss</i>	<i>0.5017</i>

Los comportamientos de estos parámetros en entrenamiento y validación se muestran en las Figuras 6 y 7 respectivamente.

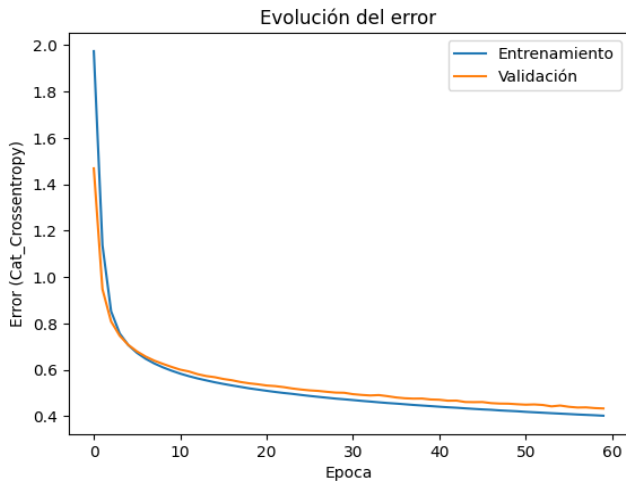


Figura 6: Gráfica de error (Categorical Cross Entropy) de la arquitectura modelo creado.

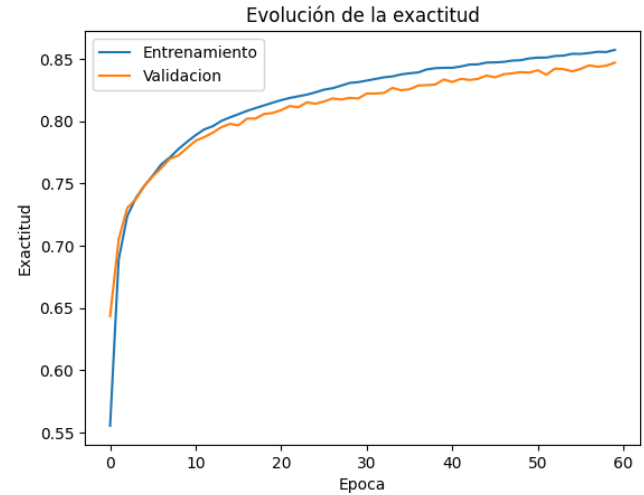


Figura 7: Gráfica de exactitud (Accuracy) de la arquitectura modelo creado.

Para la implementación se pone a prueba el modelo entrenado en un sistema de captura de vídeo en tiempo real, con el cual se pone a prueba el funcionamiento de la red detectando las posiciones de los usuarios que aparecen en la toma como se muestra en la Figura 8.

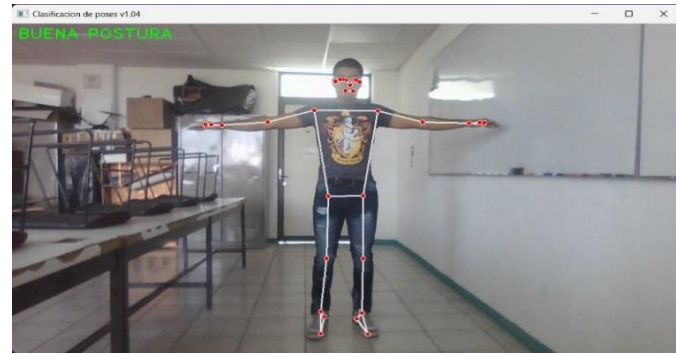


Figura 8: Prueba de implementación del funcionamiento del modelo en vídeo.

En la evaluación se realizaron 500 pruebas para verificar la correcta identificación, combinando 300 imágenes con postura correcta y 200 con postura inadecuada. Los resultados de las pruebas mostraron que de las 300 imágenes con postura correcta se clasificaron 261 como verdadero positivo y 39 con falso positivo, alcanzando un nivel de precisión en la identificación del 87% y respecto a las 200 imágenes de posturas inadecuadas se clasificaron 20 como falso positivo y 180 como verdadero negativo alcanzando un nivel de precisión en la identificación del 90% como lo muestra la matriz de confusión de la Figura 9.

	Imagen con postura correcta (300 en total)	Imagen con postura inadecuada (200 en total)
Identificación de postura correcta	261	20

identificación de postura inadecuada	39	180
--------------------------------------	----	-----

Figura 9: Matriz de confusión de las pruebas realizadas a la arquitectura.

6. Conclusiones

Los sistemas de IA aplicados al deporte han sido de gran ayuda para mejorar el rendimiento y técnicas de los deportistas, permitiendo también evitar lesiones por malas prácticas. El sistema presentado en este trabajo integra un sistema de VA, entrenado para reconocer buenas y malas posturas en patinadores profesionales, permitiendo por medio de un software de video detectar y alertar sobre las malas técnicas en la postura base.

El sistema presenta resultados adecuados en la identificación de las posturas con una red entrenada y presentada en el desarrollo de este trabajo de investigación, alcanzando porcentajes por encima del 85% en la precisión de identificación en vídeo.

El desarrollo de sistemas aplicables a diferentes disciplinas deportivas dependerá solo del establecimiento de las mecánicas, técnicas y posiciones que se desea identificar, presentando así una herramienta potencial en la asistencia de los entrenamientos y también una oportunidad de implementación tecnológica que pueda asistir a entrenadores, preparadores físicos y a los mismos deportistas que la usen buscando perfeccionar sus técnicas o mejorar su rendimiento y resultados.

Las conclusiones pueden estar al final de la discusión o en una sección separada como en este ejemplo.

Se deben reflejar las conclusiones más significativas y la importancia práctica del estudio. Las conclusiones son generalizaciones derivadas de los resultados y constituyen los aportes y las innovaciones del estudio realizado.

7. Referencias

Annino, G., Bonaiuto, V., Campoli, F., Caprioli, L., Edriss, S., Padua, E., ... & Zanela, A. (2023, September). Assessing sports performances using an artificial intelligence-driven system. In *2023 IEEE International Workshop on Sport, Technology and Research (STAR)* (pp. 98-103). IEEE.

Bin Li, X. X. (2021). Application of Artificial Intelligence in Basketball Sport. *Journal of Education, Health and Sport*, 14.

Bodemer, O. (2023). Enhancing individual sports training through artificial intelligence: A comprehensive review. *Authorea Preprints*.

Chen, Y., Zhu, Y., Papandreou, G., & Yuille, A. (2017). Associative embedding. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4107-4115).

Chunguang Li, J. C. (2023). Retracted: Intelligent Sports Training System Based on Artificial. *Hindawai*, 12.

Cruz, M. d. (Julio de 14 de 2021). La visión por computadora y las futuras aplicaciones tecnológicas en diversos escenarios. Obtenido de ESPE: <https://journal.espe.edu.ec/ojs/index.php/Academia-de-guerra/article/view/VOL12ART13>

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.

He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 2961-2969).

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-excitation networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 7132-7141).

J. Riera, N., P. O., R. N. Verazay, A., Paz, F., Battezzati, V., Chuca, R., . . . Arjona, F. (2020). Técnicas de inteligencia artificial aplicadas a problemas de visión por computadora. *Semantic Scholar*.

Kasper, K. (2019). Sports training principles. *Current sports medicine reports*, 18(4), 95-96.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 1097-1105.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (2016). SSD: Single shot multibox detector. In *European Conference on Computer Vision* (pp. 21-37). Springer.

Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3431-3440).

Luckstead, E. F., Satran, A. L., & Patel, D. R. (2002). Sport injury profiles, training and rehabilitation issues in American sports. *Pediatric Clinics*, 49(4), 753-767.

Nadikattu, R. R. (2020). Implementation of new ways of artificial intelligence in sports. *Journal of Xidian University*, 14(5), 5983-5997.

Naik, B. T., Hashmi, M. F., & Bokde, N. D. (2022). A comprehensive review of computer vision in sports: Open issues, future trends and research directions. *Applied Sciences*, 12(9), 4429.

Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.

Redmon, J., Farhadi, S., & others. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).

Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems*, 28, 91-99.

Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556**.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition.

Viru, A. (2017). *Adaptation in sports training*. Routledge.

Wang, Y., & Yu, Z. (2020). Deep learning for image super-resolution: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(4), 789-809.

Zhang, R., Isola, P., Efros, A. A., & Heeger, D. J. (2016). Colorful image colorization. In *European Conference on Computer Vision* (pp. 649-666). Springer.