

INSTITUTO TECNOLÓGICO DE CHIHUAHUA
DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN

***“CIENCIA DE DATOS PARA LA SALUD Y
PROCESAMIENTO DE SEÑALES DE
MOVIMIENTO DE MIEMBROS SUPERIORES”***

TESIS

QUE PARA OBTENER EL GRADO DE

***MAESTRA EN CIENCIAS EN INGENIERÍA
ELECTRÓNICA***

PRESENTA:

SAHYRA JANETH BORUNDA VALENZUELA

**DIRECTOR DE LA TESIS:
DR. JUAN ALBERTO RAMÍREZ QUINTANA**

CHIHUAHUA, CHIH., ENERO 2025



Educación
Secretaría de Educación Pública



TECNOLÓGICO
NACIONAL DE MÉXICO®





Educación
Secretaría de Educación Pública



TECNOLÓGICO
NACIONAL DE MÉXICO



Instituto Tecnológico de Chihuahua
División de Estudios de Posgrado e Investigación

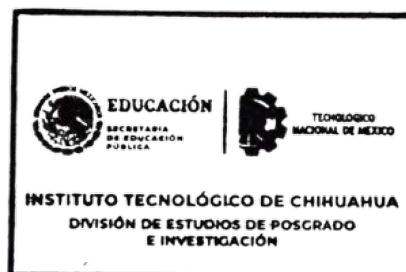
Chihuahua, Chih. **21 de enero de 2025**

**C. SAHYRA JANETH BORUNDA VALENZUELA
PRESENTE**

En cumplimiento con los requerimientos para la obtención del grado de Maestra en Ciencias en Ingeniería Electrónica y a propuesta de su Comité Tutorial, la División de Estudios de Posgrado e Investigación le concede la autorización para imprimir la tesis titulada **"Ciencia de datos para la salud y procesamiento de señales de movimiento de miembros superiores"** dirigida por el Dr. Juan Alberto Ramírez Quintana con el siguiente contenido de capítulos:

- I Marco teórico del M²S
- II Antecedentes
- III Análisis de datos del M²S
- IV Flujo óptico para análisis de M²S
- V Redes neuronales para clasificación de M²S
- VI Conclusiones

ATENTAMENTE
Excelencia en Educación Tecnológica®
"La Técnica por el Engrandecimiento de México"



LEONCIO OCHOA CERVANTES
JEFE DE LA DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN

LOC/adcs



2025
Año de
**La Mujer
Indígena**

Av. Tecnológico #2909 Col. 10 de Mayo
C.P. 31200, Chihuahua, Chihuahua. Tel. (614) 2012000 ext. 2157
e-mail: dir_chihuahua@tecnm.mx tecnm.mx <http://itchihuahua.mx/>





Chihuahua, Chih. **20 de enero de 2025**

LEONCIO OCHOA CERVANTES
JEFE DE LA DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN
PRESENTE

En cumplimiento con los requerimientos para la obtención del grado de Maestra en Ciencias en Ingeniería Electrónica, le notificamos que el documento de tesis de la alumna **C. SAHYRA JANETH BORUNDA VALENZUELA**, titulado **"Ciencia de datos para la salud y procesamiento de señales de movimiento de miembros superiores"** dirigido por el Dr. Juan Alberto Ramírez Quintana, ha sido aprobado y aceptado para su impresión.

Por lo anterior, proponemos le sea concedida la autorización de impresión correspondiente.

Agradeciendo la atención a la presente, quedamos de usted:

ATENTAMENTE

*Excelencia en Educación Tecnológica
La técnica por el Engrandecimiento de México*

Juan Ramirez

DR. JUAN ALBERTO RAMÍREZ QUINTANA
MIEMBRO DEL COMITÉ TUTORIAL

DR. JESÚS ALFONSO MEDRANO
HÉRMOSILLO
MIEMBRO DEL COMITÉ TUTORIAL

Graciela Ramirez A

DRA. GRACIELA MARÍA DE JESÚS RAMÍREZ
ALONSO
MIEMBRO DEL COMITÉ TUTORIAL

MTRA. ALMA DELIA CORRAL SÁENZ
MIEMBRO DEL COMITÉ TUTORIAL



2025
Año de
La Mujer
Indígena

Av. Tecnológico #2909 Col. 10 de Mayo
C.P. 31200, Chihuahua, Chihuahua. Tel. (614) 2012000 ext. 2157
e-mail: dir_chihuahua@tecnm.mx <http://itichihuahua.mx/>





Educación
Secretaría de Educación Pública



TECNOLÓGICO
NACIONAL DE MÉXICO



Instituto Tecnológico de Chihuahua
División de Estudios de Posgrado e Investigación

CARTA CESIÓN DE DERECHOS

En la ciudad de Chihuahua el día 21 de enero de 2025, el que suscribe, C. SAHYRA JANETH BORUNDA VALENZUELA con número de control G22061171, de la Maestría en Ciencias en Ingeniería Electrónica, adscrita a la División de Estudios de Posgrado e Investigación del Instituto Tecnológico de Chihuahua, manifiesta que es autor intelectual del presente trabajo de Tesis bajo la dirección del el Dr. Juan Alberto Ramírez Quintana y cede los derechos del trabajo titulado "Ciencia de datos para la salud y procesamiento de señales de movimiento de miembros superiores", al Tecnológico Nacional de México y/o Instituto Tecnológico de Chihuahua para su difusión, divulgación, transmisión, reproducción, así como su digitalización con fines académicos y de investigación.

C. SAHYRA JANETH BORUNDA VALENZUELA



2025
Año de
La Mujer
Indígena

Av. Tecnológico #2909 Col. 10 de Mayo
C.P. 31200, Chihuahua, Chihuahua. Tel. (614) 2012000 ext. 2157
e-mail: dir_chihuahua@tecnm.mx tecnm.mx <http://itchihuahua.mx>





Educación
Secretaría de Educación Pública



TECNOLÓGICO
NACIONAL DE MÉXICO



Instituto Tecnológico de Chihuahua
División de Estudios de Posgrado e Investigación

DECLARACIÓN DE ORIGINALIDAD

En la ciudad de Chihuahua el día 21 de enero de 2025, el que suscribe, C. SAHYRA JANETH BORUNDA VALENZUELA de la Maestría en Ciencias en Ingeniería Electrónica, con número de control G22061171, adscrito a la División de Estudios de Posgrado e Investigación del Instituto Tecnológico de Chihuahua, manifiesta que es autor intelectual de la tesis titulada "Ciencia de datos para la salud y procesamiento de señales de movimiento de miembros superiores" bajo la dirección el Dr. Juan Alberto Ramírez Quintana; que el contenido es original y que las fuentes de información consultadas para su fundamentación están debidamente citadas y referenciadas.

C. SAHYRA JANETH BORUNDA VALENZUELA



2025
Año de
La Mujer
Indígena

Av. Tecnológico #2909 Col. 10 de Mayo
C.P. 31200, Chihuahua, Chihuahua. Tel. (614) 2012000 ext. 2157
e-mail: dir_chihuahua@tecnm.mx tecnm.mx <http://itchihuahua.mx>

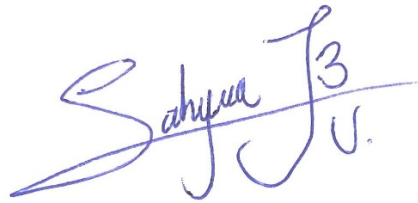


Chihuahua, Chihuahua, 24 de enero de 2021

Dra. María Elena Álvarez-Buylla Roces
Director/a de Conahcyt

Sirva la presente para saludarla e informarle que a la fecha he obtenido el grado de **Maestro en Ciencias en Ingeniería Electrónica** en la División de Estudios de Posgrado e Investigación del Instituto Tecnológico de Chihuahua, motivo por el cual agradezco todo el apoyo brindado por este Consejo que usted representa, con el otorgamiento de una beca que permitió dedicarme de manera exclusiva a la realización de mis estudios de posgrado y de esta manera lograr el objetivo principal del convenio establecido.

Sin otro particular quedo a sus órdenes no sin antes reiterar mi agradecimiento.
Atentamente

A handwritten signature in blue ink, appearing to read 'Sahyra J. B. V.' with a stylized flourish.

Sahyra Janeth Borunda Vaenzuela

AGRADECIMIENTOS

Agradezco al Tecnológico Nacional de México por el apoyo brindado con el proyecto 19182.24-P.

Quiero expresar mi más profundo agradecimiento a todas aquellas personas que de alguna manera han sido mi soporte y han formado parte de esta travesía académica.

A mis padres, Olivia y Juan, que con su infinito amor y apoyo han confiado en mí a cada paso que doy, me han dado la confianza para ser yo misma y buscar mi camino a la felicidad. Gracias a ustedes he formado mi valor y carácter que me ha llevado hasta donde estoy y mucho más. Agradezco el soporte que me dan a cada momento, me han ofrecido las herramientas para alcanzar mis metas, además de su comprensión y cariño cuando mas lo he necesitado. Ustedes lo son todo.

Agradezco a mi hermana Estefany, por siempre estar y ser un refuerzo en cada paso que he dado a lo largo de mi vida. A mis sobrinos, Rodolfo y Rodrigo son la luz de cada día, me impulsan a ser mejor y me llevan al conocimiento con cada una de sus preguntas y respuestas.

Gracias a mi asesor, por su orientación y apoyo constante e incondicional, por guiarme en cada paso de este trayecto. Gracias por tener siempre la disposición de ayudar en cada inquietud existente desde el momento en que inicie este recorrido.

Doy gracias a mis amigos, que me apoyaron e incitaron a buscar el camino del conocimiento, que han creído en mí y me han acompañado todo este tiempo.

Agradezco a las personas que ya no están, pero las llevo siempre en el pensamiento. A pesar de la distancia, siempre estaremos juntos.

Este logro es de todos los que me han acompañado.

RESUMEN

CIENCIA DE DATOS PARA PROCESAMIENTO DE SEÑALES DE MOVIMIENTO DE MIEMBROS SUPERIORES

Sahyra Janeth Borunda Valenzuela

Maestría en Ciencias en Ingeniería Electrónica

División de Estudios de Posgrado e Investigación del

Instituto Tecnológico de Chihuahua

Chihuahua, Chih. 2024

Director de Tesis: Dr. Juan Alberto Ramírez Quintana

Los miembros superiores son partes del cuerpo humano que realizan funciones relacionadas con la manipulación de objetos y se componen del hombro, brazo y antebrazo. Conocer el funcionamiento de estas partes del cuerpo es de vital importancia para el análisis del movimiento humano y para el seguimiento de lesiones deportivas. Existen diferentes técnicas para conocer mejor los miembros superiores, entre las que destacan el uso de sensores magnético-inerciales y cámaras calibradas para estimación de movimiento. Sin embargo, este tema ha sido poco explorado en la literatura científica y la tecnología. Por lo cual, en esta tesis se propone un esquema de ciencia de datos para el análisis del movimiento de miembros superiores durante el lanzamiento de un balón de fútbol americano. Este esquema inició con la creación de la base de datos DPMov (*Direct Pass Movement Dataset*) que contiene 106 videos de voluntarios ejecutando un lanzamiento de pase directo con un balón de fútbol. Cada video se realizó aplicando un protocolo de adquisición y se registraba la información del voluntario o voluntaria como peso, estatura, edad, si jugaba fútbol, etc. A partir del registro de información, se realizó un análisis categórico para definir las clases del modelo de ciencia de datos con los algoritmos de *K-Means* y *Fuzzy C-Means* para diferenciar entre personas que tienen una técnica de lanzamiento adecuada y las que no. Por otro lado, los videos se utilizaron como entrada de un modelo de aprendizaje automático que analiza y clasifica el patrón del lanzamiento para predecir si este fue realizado por un jugador frecuente o infrecuente. El

modelo de aprendizaje automático consta de una etapa flujo óptico denso llamado *Gunnar Farneback*, para obtener estadísticas de la media, mediana, entropía y curtosis de la dirección del movimiento durante el lanzamiento del balón y una Red Neuronal Artificial (RNA) para analizar dichas estadísticas. Los experimentos consistieron en comparar el desempeño con diferentes RNA con la base de datos de DPMov. Los resultados mostraron que la red recurrente con neuronas de memoria de corto y largo plazo tuvo una precisión (*accuracy*) de 0.81, la cual obtuvo el mejor desempeño para detectar la experiencia de un jugador y el riesgo de lesión de miembros superiores.

CONTENIDO

INTRODUCCIÓN	i
I. MARCO TEÓRICO DEL M²S.....	1
1.1 Conceptos de miembros superiores.....	1
1.1.1 Estructura ósea de miembros superiores.....	1
1.1.2 Estructura muscular de miembros superiores.....	2
1.2 Lesiones en miembros superiores	6
1.2.1 Lesiones de hombro	6
1.2.2 Lesiones del codo	9
1.2.3 Lesiones de muñeca y mano.....	13
1.3 Rendimiento deportivo de miembros superiores.....	15
II. ANTECEDENTES	17
2.1 Flujo Óptico	17
2.1.1 Flujo óptico disperso	18
2.1.2 Flujo óptico denso	19
2.2 Inteligencia artificial	19
2.2.1 Ciencia de datos	19
2.2.2 Aprendizaje de maquina.....	20
2.2.3 Aprendizaje profundo.....	21
2.3 Redes neuronales Artificiales.....	21
2.4 Análisis del estado del arte.....	24
2.4.1 Técnicas de medición de movimiento humano	25
2.4.2 Ciencia de datos para análisis de movimiento humano.....	25
III. ANÁLISIS DE DATOS DEL M²S.....	38
3.1 Protocolo	38
3.1.1. Escenario	39
3.1.2 Toma de datos de voluntarias y voluntarios	40
3.2 Descripción de la base de datos.....	41
3.2.1 Lanzamiento directo de balón	42
3.2.2 Lanzamiento del balón en voluntarios.....	45
3.3 Análisis categórico	46
3.3.1 Coeficiente de correlación de <i>Pearson</i>	49

3.3.2	Análisis de Componentes Principales	50
3.3.3	<i>Clustering</i> no Supervisado con <i>K-Means</i>	53
3.3.4	<i>Fuzzy C-Means</i> (FCM).....	55
3.4	Resultados del análisis categórico.....	56
3.4.1	Datos estadísticos de jugador o no jugador	57
3.4.2	Resultados de K-Means.....	57
3.4.3	Resultados de FCM	58
IV.	FLUJO ÓPTICO PARA EL ANÁLISIS DE M²S.....	61
4.1	Flujo óptico de <i>Gunnar Farneback</i>	61
4.2	Cálculo del flujo óptico	63
4.3	Estadísticas	67
4.3	Dimensión de datos	71
V.	REDES NEURONALES PARA CLASIFICACIÓN DE M²S.....	75
5.1	Perceptrón multicapa.....	75
5.1.1	Aprendizaje de la MLP	77
5.1.2	Arquitectura de la red.....	78
5.2	Redes neuronales convolucionales (CNN).....	79
5.2.1	Arquitectura de la red.....	80
5.3	Redes neuronales recurrentes	81
5.3.1	RNN simple.....	81
5.3.2	Memorias de corto y largo plazo.....	82
5.4	Resultados	83
5.4.1	Parámetros de aprendizaje.....	83
5.4.2	Resultados de las métricas.....	83
VI.	CONCLUSIONES.....	85
VII.	REFERENCIAS.....	88

INTRODUCCIÓN

El análisis del movimiento del cuerpo humano es de vital importancia en una amplia gama de disciplinas como la medicina, la rehabilitación, el deporte, la ergonomía y la robótica. Esta forma de análisis involucra la interpretación del movimiento de la pose, expresiones faciales, miembros inferiores y miembros superiores. El análisis de pose y expresiones faciales es útil para conocer el estado emocional de una persona y en la prevención de lesiones en cuello y vertebras. El análisis de miembros inferiores proporciona una visión detallada y cuantitativa de cómo se mueven las piernas y cómo interactúan con el entorno [1]. Por otro lado, el análisis de los miembros superiores tiene las siguientes aplicaciones:

- Identificar condiciones como lesiones deportivas, síndrome del túnel carpiano, o hemiparesia tras un accidente cerebrovascular.
- Monitorear el progreso de la recuperación de los pacientes tras una lesión o cirugía, ajustando los planes de tratamiento de acuerdo con las necesidades individuales.
- Rehabilitación de accidentes que afecten los miembros superiores, como fracturas o roturas de ligamentos.
- Perfeccionar la técnica, mejorar el rendimiento y prevenir lesiones en deportes como el Fútbol Americano, Tenis, Béisbol o Natación.
- Ayuda a comprender cómo se coordinan los músculos, tendones y articulaciones para realizar movimientos complejos. Esta comprensión es crucial en el desarrollo de tecnologías que imiten o asistan el movimiento, como robots colaborativos o prótesis inteligentes.

Respecto a la literatura científica, el seguimiento del movimiento corporal ha sido objeto de gran interés desde la década de 1990, por lo que se han reportado diversas técnicas de medición, como sistemas ópticos, mecánicos, magnético-inerciales, entre otros [2]. Además, en años recientes se han presentado también trabajos relacionados a de ciencia de datos para tareas como como reconocimiento de actividades y deportes [2].

La ciencia de datos ha jugado un papel crucial en el análisis del movimiento humano, con métodos que incluyen el uso de sensores magnético-inerciales para estimar la pose corporal en

INTRODUCCIÓN

tiempo real y la combinación de técnicas de visión por computadora para el reconocimiento de acciones humanas en deportes y entornos industriales [3] [4]. Aunado a esto, en [5] se destaca en la importancia del aprendizaje automático y las redes neuronales en la mejora del análisis de movimiento humano y en la construcción de sistemas de reconocimiento más precisos, recalcando avances en la eficiencia de la captura de movimientos en tiempo real y la clasificación de actividades.

Sin embargo, debido a la complejidad biomecánica y limitaciones tecnológicas el tema del análisis de miembros superiores ha sido menos explorado en comparación con otros aspectos del análisis del movimiento por varias razones clave como:

1. Complejidad anatómica y funcional. Son partes del cuerpo anatómicamente complejos, lo que causa lo siguiente:

Mayor complejidad del movimiento. Los miembros superiores tienen una anatomía y una gama de movimientos mucho más compleja que los miembros inferiores. Las articulaciones del hombro, codo y muñeca permiten movimientos en múltiples planos, lo que dificulta la medición y análisis preciso de las fuerzas y tensiones que generan lesiones.

Diversidad de movimientos finos y gruesos. Las manos y brazos son responsables de movimientos tanto gruesos (como levantar o empujar) como finos (escribir, manipular objetos pequeños), lo que complica el análisis unificado.

2. Menor incidencia en estudios clínicos. Se realizan menos estudios clínicos en miembros superiores por las siguientes razones:

Foco en las extremidades inferiores: Tradicionalmente, se ha dado más énfasis a las extremidades inferiores debido a la alta incidencia de problemas relacionados con la marcha, la movilidad, y la importancia de las piernas en tareas diarias como caminar, correr o mantener el equilibrio.

Lesiones más críticas en miembros inferiores: Las lesiones en los miembros inferiores pueden tener un impacto más severo en la movilidad y calidad de vida, lo que ha motivado una mayor cantidad de estudios e investigaciones.

3. Desafíos tecnológicos. Debido a la complejidad y pocos estudios, es menor el desarrollo tecnológico para el análisis de miembros superiores que para otras partes del cuerpo. Esto causa lo siguiente:

Dificultades para medir movimientos complejos: los métodos de captura de movimiento como los sistemas de seguimiento óptico o los sensores inerciales son más fáciles de aplicar a movimientos lineales o repetitivos (comunes en las piernas). Capturar y analizar movimientos complejos de los brazos o las manos requiere tecnologías más sofisticadas y precisas, lo que ha ralentizado la investigación en este ámbito.

Falta de estandarización: aún no existen protocolos estandarizados y universalmente aceptados para evaluar lesiones en los miembros superiores, lo que limita la cantidad de estudios comparables.

En conjunto, estos factores han contribuido a que el tema del movimiento de miembros superiores no haya sido tan explorado como otros temas de movimiento corporal. Sin embargo, en los últimos años ha habido un creciente interés debido al desarrollo de nuevas tecnologías y una mayor conciencia sobre la importancia de prevenir lesiones en ámbitos como el deporte y la ergonomía laboral. Además, la evaluación precisa y cuantitativa del movimiento de los miembros superiores permite diseñar programas de rehabilitación personalizados para jugadores que han sufrido lesiones [6].

Para contribuir con este esfuerzo, el objetivo de esta tesis es desarrollar un esquema de ciencia de datos para el análisis del movimiento de los miembros superiores durante el lanzamiento de un pase directo realizado típicamente en deportes como el Fútbol Americano. Este esquema incluye lo siguiente:

- Base de datos DPMov (*Direct Pass Movement Dataset*) que contiene 106 videos de voluntarios ejecutando un lanzamiento de pase directo con un balón de Fútbol.
- Análisis categórico para definir clases que agrupan personas con experiencia en realizar el pase directo y personas que no tienen experiencia.
- Modelo de aprendizaje automático que predice las clases definidas en el análisis categórico.

INTRODUCCIÓN

El movimiento de pase directo realizado en el Fútbol Americano se seleccionó en esta tesis porque permite analizar el movimiento de los brazos y hombros y puede ayudar a prevenir lesiones comunes en los jugadores de Fútbol Americano, como desgarros de ligamentos en el hombro o tendinitis.

El modelo de ciencia de datos de esta tesis se divide en dos etapas: la generación y análisis de datos y el algoritmo de aprendizaje automático que interpreta el movimiento.

1. Generación y análisis de datos

La Figura 1 presenta un esquema general de la generación y análisis de datos. La parte de generación consiste en la creación de la base de datos desarrollada para esta tesis. Dado que no existían suficientes datos sobre el movimiento de miembros superiores en el contexto del Fútbol Americano, se diseñó la base de datos DPMov (*Direct Pass Movement Dataset*) para llenar este vacío. Esta base de datos busca capturar y analizar el movimiento de lanzamiento directo de balón de Fútbol Americano, permitiendo una mejor visualización y comprensión de la técnica.

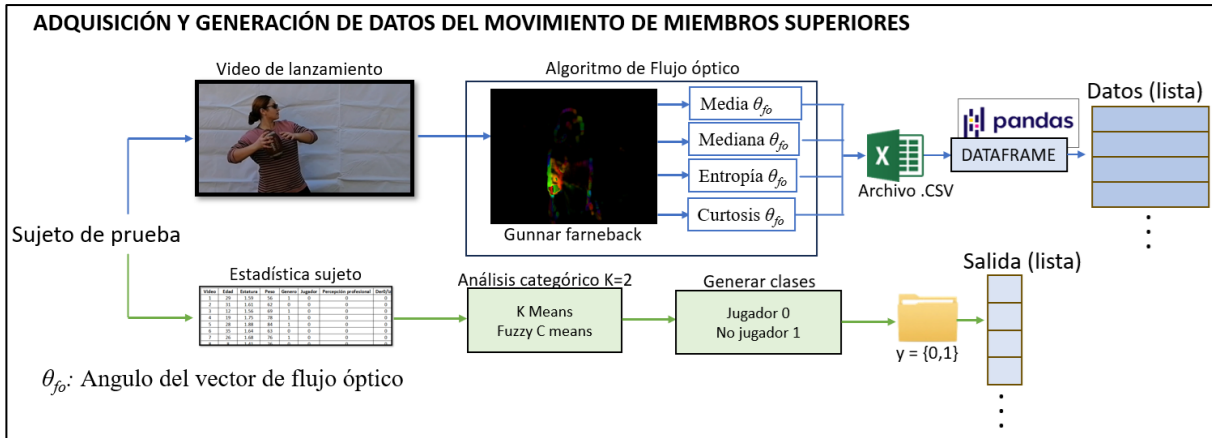


Figura 1. Esquema general para adquirir y analizar datos de miembros superiores.

La captura del movimiento de balón se realizó con un video y un protocolo que incluyó la grabación de videos con un dispositivo móvil, un trípode para garantizar uniformidad en las tomas y una lona blanca como fondo. Los participantes fueron jugadores frecuentes e infrecuentes que proporcionaron datos antropométricos como estatura, peso, y si son jugadores con experiencia o no.

INTRODUCCIÓN

El análisis de datos tiene dos fases: flujo óptico y análisis categórico. El flujo óptico permite medir el movimiento de los brazos y manos durante el lanzamiento de un balón. Se utilizó el algoritmo de *Gunnar Farneback* para generar mapas de movimiento donde los vectores de desplazamiento indican la dirección y magnitud del movimiento en diferentes puntos de los brazos. Posteriormente, se obtienen características de media, mediana, entropía y curtosis de las direcciones definidas por el flujo óptico. Estas características se utilizarán como entrada para entrenar y probar una red neuronal. El análisis categórico utilizó técnicas estadísticas no supervisadas para identificar y generar clases jugadores frecuentes e infrecuentes basados en variables como edad, estatura, peso, sexo y características biomecánicas.

Estos datos se utilizaron para modelar la técnica de lanzamiento y explorar diferencias entre grupos, proporcionando información valiosa para futuras investigaciones en el deporte.

2. Algoritmo de procesamiento

El algoritmo de procesamiento se muestra en la Figura 2 y tiene como entrada los datos de media, mediana, entropía y curtosis obtenidos en cada video para clasificar si el sujeto tiene experiencia lanzando un balón de futbol americano o no. Las etiquetas para la clasificación se basan en que cada sujeto indica si tiene experiencia lanzando (Jugador) o no (No jugador).

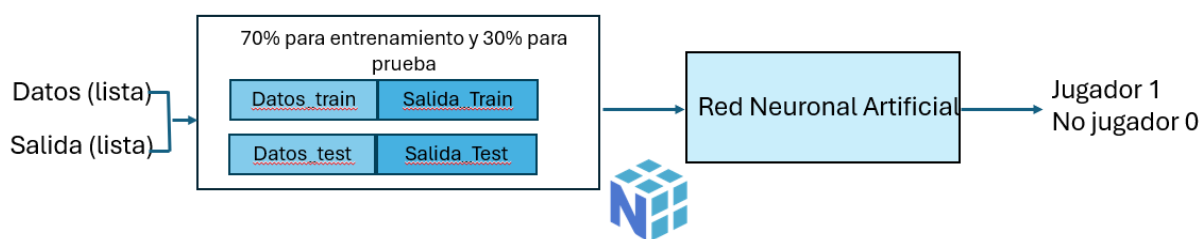


Figura 2. Algoritmo de procesamiento.

El algoritmo de procesamiento se implementó en Python y realiza los siguientes pasos:

Los datos de las características se dividieron de la siguiente manera:

- El 70% de los datos se utilizó para entrenar el modelo. Este conjunto de entrenamiento permitió que el algoritmo aprendiera las diferencias y patrones en los movimientos de jugadores frecuentes y no frecuentes.

INTRODUCCIÓN

- El 30% restante de los datos se reservó para pruebas, con el fin de validar la precisión del modelo y verificar su capacidad para generalizar a nuevos casos.

Se utilizó un modelo de clasificación supervisado que tomó como entrada las métricas obtenidas (entropía, curtosis, media, mediana). El algoritmo se basó en una red neuronal artificial con arquitectura Perceptrón Multicapa (MLP por sus siglas en inglés), Redes Neuronales Convolucionales (CNN por sus siglas en inglés) y Redes Neuronales Recurrentes (RNN por sus siglas en inglés). El objetivo del entrenamiento fue identificar los patrones subyacentes que separan a los jugadores frecuentes (con experiencia) de los no jugadores o jugadores inexpertos.

Durante el entrenamiento, las redes neuronales aprendieron a asociar las estadísticas del flujo óptico con el nivel de habilidad de los participantes, lo que permitió al modelo "entender" las características biomecánicas de un lanzamiento eficiente y bien entrenado frente a uno menos desarrollado. Los resultados de la prueba fueron comparados con la clasificación real de los participantes, lo que permitió evaluar la precisión del modelo y hacer ajustes si era necesario para mejorar su capacidad de predicción.

3. Descripción de la tesis

El resto de la tesis se organiza de la siguiente manera: el primer capítulo es el marco teórico del Movimiento de Miembros Superiores (M²S), donde se describe la anatomía y el funcionamiento de las estructuras que conforman los miembros superiores, así como las lesiones asociadas al movimiento y los indicadores del rendimiento deportivo. El segundo capítulo es sobre los antecedentes, donde se describen los métodos que se utilizaron para el desarrollo de esta tesis, tales como el flujo óptico, aprendizaje de máquina, aprendizaje profundo y redes neuronales artificiales. El tercer capítulo trata sobre el análisis de datos de M²S, donde se describe el protocolo para la elaboración de la base de datos, el análisis categórico basado en agrupamiento y sus resultados. El cuarto capítulo habla sobre el flujo óptico para el análisis en el M²S, donde se presenta el método de *Gunnar Farnebäck* y su análisis. A partir de esto, se presentan las estadísticas de las direcciones de flujo óptico que se utilizaran para el análisis de M²S. El capítulo 5 trata del algoritmo de procesamiento basado en

INTRODUCCIÓN

redes neuronales artificiales que interpretan y clasifica el M²S en jugador y no jugador. El capítulo 6 es una discusión de los resultados obtenidos por la tesis y las conclusiones. Las conclusiones presentan un conjunto de hipótesis de cómo se relaciona la anatomía y padecimientos de miembros superiores con las estadísticas de flujo óptico y las redes neuronales.

El termino Movimiento de Miembros Superiores se referirá como M²S durante el resto de la tesis para simplificar el cuerpo del texto del presente documento.

I. MARCO TEÓRICO DEL M²S

Los miembros superiores son las extremidades del cuerpo humano que se extienden desde el tronco y se componen de hombro, brazo y mano. Estas extremidades desempeñan un papel esencial en la vida cotidiana de los seres humanos, ya que están diseñados para la manipulación de objetos mediante el movimiento fino y complejo. El conocimiento profundo de cómo se mueve y funciona el tracto superior es fundamental para comprender y abordar una amplia variedad de desafíos. Estos desafíos incluyen la rehabilitación de pacientes con lesiones o discapacidades en los miembros superiores, la mejora de la eficiencia de las tareas laborales, la optimización de la ergonomía en entornos de trabajo y el diseño de dispositivos y tecnologías avanzadas.

Para tener un mejor conocimiento del M²S y su función en el tracto superior, en este capítulo se examinará la anatomía, funcionamiento y biomecánica de miembros superiores. La sección 1.1 presenta los conceptos básicos de miembros superiores. La sección 1.2 describe las lesiones asociadas a los brazos, hombros y mano. Finalmente, la sección 1.3 presenta los indicadores de rendimiento deportivo.

Para comprender mejor los términos de este capítulo se puede consultar el libro [7] sección XI.

1.1 Conceptos de miembros superiores

La morfología del miembro superior depende de su estructura ósea, muscular y se divide en brazo, antebrazo, mano y dedos. Estos elementos se definen a continuación.

1.1.1 Estructura ósea de miembros superiores

La estructura ósea (también conocida como esqueleto), es el marco interno que proporciona soporte, protección y forma al cuerpo. Consta de huesos, cartílagos y ligamentos y se divide en el esqueleto axial y el apendicular. El esqueleto axial son los huesos que brindan soporte al cuerpo, mientras que el apendicular da forma a la cadera, brazos y piernas.

I. MARCO TEÓRICO DEL M²S

La estructura ósea de los miembros superiores forma parte del esqueleto apendicular e inicia en los huesos de las escapulas y la clavícula, los cuales conectan el brazo con el tronco. El brazo tiene tres segmentos de huesos como se muestra en la Figura 1.1. El primer segmento del brazo se constituye por un solo hueso, el húmero. El segundo segmento o antebrazo comprende el cúbito y el radio. El tercer segmento se compone de muñeca o carpo, mano o metacarpo y dedos o falanges.

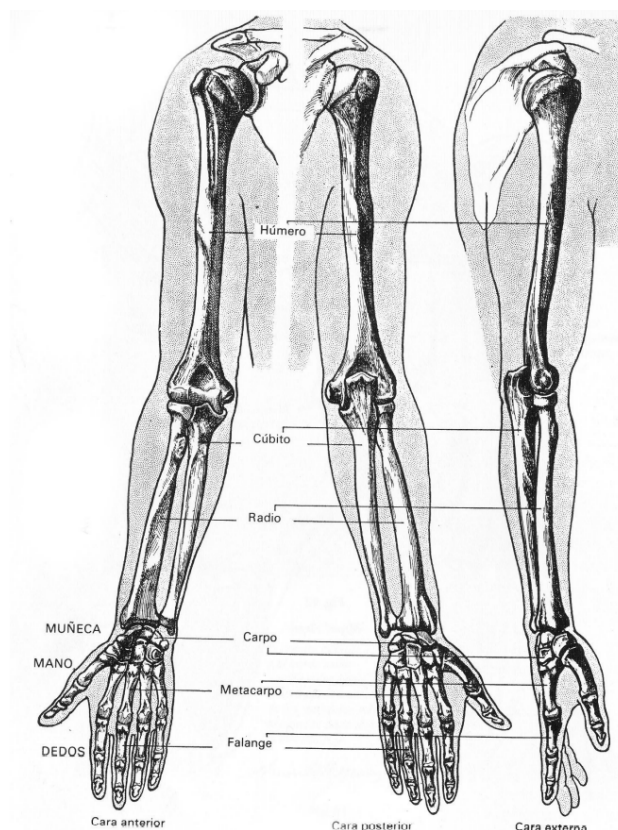


Figura 1.1. Esquema osteológico general del miembro superior Imagen tomada de [8].

1.1.2 Estructura muscular de miembros superiores

La estructura muscular se refiere a la organización de los músculos, los cuales son tejidos blandos que pueden contraerse y retraerse, permitiendo el movimiento del cuerpo humano. Existen tres tipos de músculos: esqueléticos, cardíacos y lisos. Los músculos esqueléticos son tejidos voluntarios (controlados conscientemente) y están conectados a los huesos. Los músculos cardíacos son tejidos involuntarios (no controlados) que se encargan de bombear la

sangre a través del sistema circulatorio. Los músculos lisos son tejidos involuntarios que controlan el funcionamiento de diversos órganos. Entre estos, los músculos de miembros superiores son susceptibles a lesiones son los esqueléticos, ya que se pueden mover de forma voluntaria en actividades de alto rendimiento que sobrepasen las capacidades de dichos tendones. Estos músculos forman estructuras corporales que se describen a continuación.

Hombro. El hombro está conformado por el conjunto muscular que va de la escápula al húmero. Estos son el deltoides, el subescapular, el supraespinoso, el infraespinoso, el redondo mayor y el redondo menor [9]. En la Figura 1.2 se presenta su composición muscular.

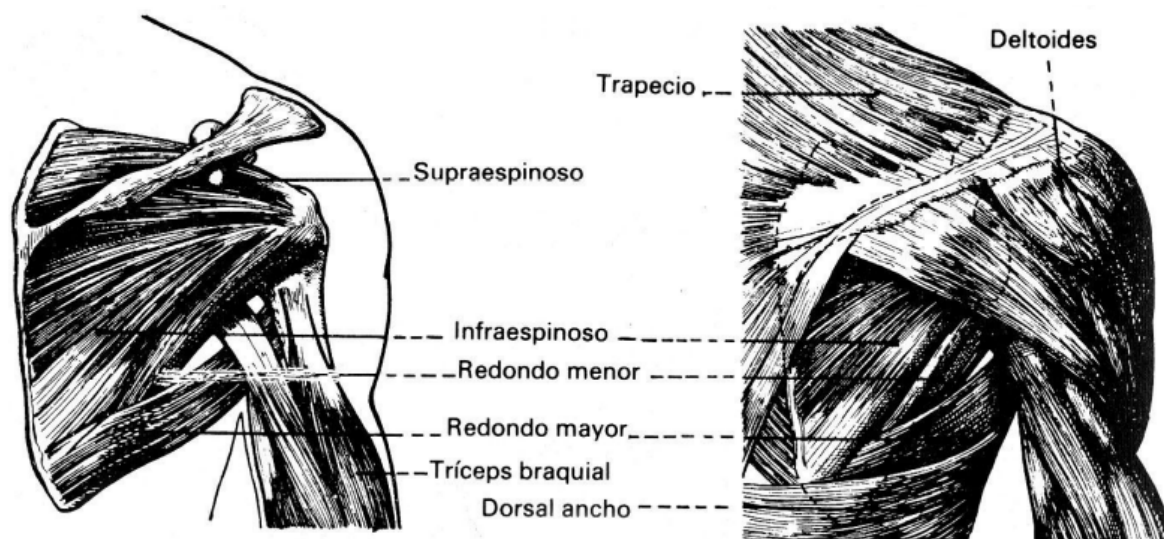


Figura 1.2. Composición muscular del hombro. Imagen tomada de [8]

Brazo. Se constituye de cuatro músculos: bíceps que están en la región anterior, tríceps que está en la región posterior, el braquial y el coracobraquial. En la Figura 1.3 se observa la diferenciación muscular del brazo.

Antebrazo. Los músculos del antebrazo son un grupo de 20 músculos distribuidos en región anterior, posterior y externa. La región anterior se compone de ocho músculos: cuatro superficiales y cuatro profundos. Los músculos profundos están compuestos por: flexor profundo de los dedos, flexor largo del pulgar, pronador cuadrado. Los músculos superficiales son el pronador redondo, el palmar mayor, el palmar menor, el cubital anterior más el flexor

I. MARCO TEÓRICO DEL M²S

superficial de los dedos. La región externa se compone de supinador largo, primer radial externo, segundo radial externo y supinador corto.

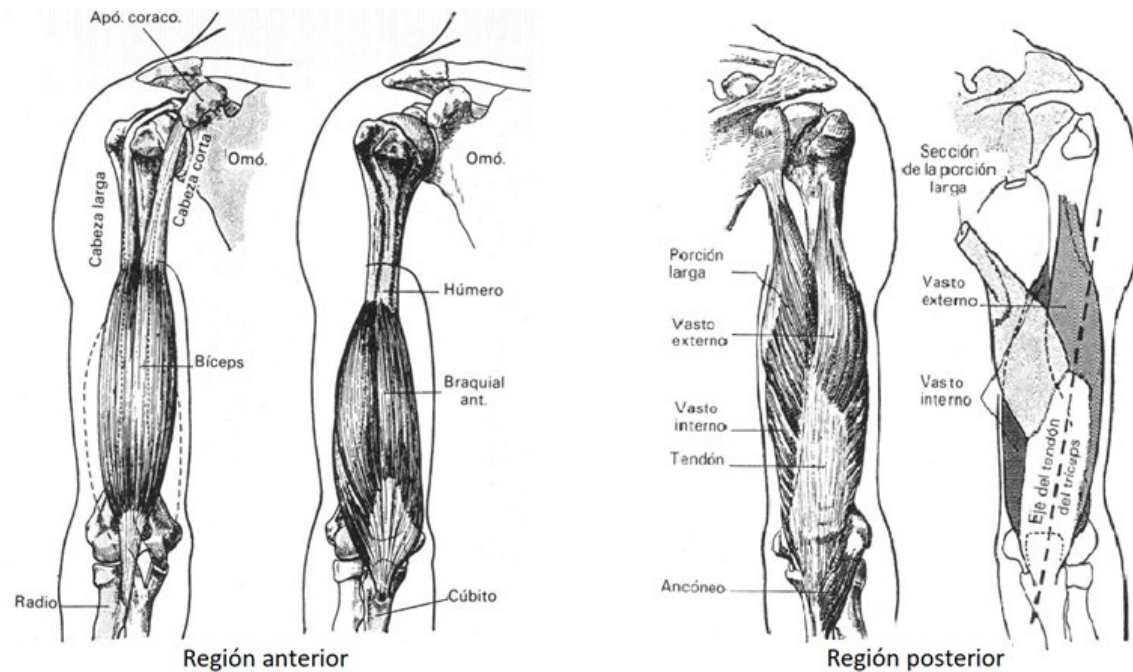


Figura 1.3. Composición muscular del brazo. Imagen tomada de [8].

La región posterior dispone de ocho músculos colocados en dos capas superpuestas. La capa superficial se compone de cuatro músculos cuyo modelado es superficial en los dos tercios superiores del antebrazo. Está formada por el ancóneo, el cubital superior, el extensor del meñique y el extensor común. Todos tienen un origen en un tendón común y luego se individualizan. La capa profunda está formada por cuatro músculos cuyo modelado se realiza en el tercio inferior del antebrazo y en la muñeca. Se conforma con el abductor largo, extensor corto, extensor largo del pulgar, y extensor del índice. En la Figura 1.4 se percibe a través de diferentes caras la disposición muscular.

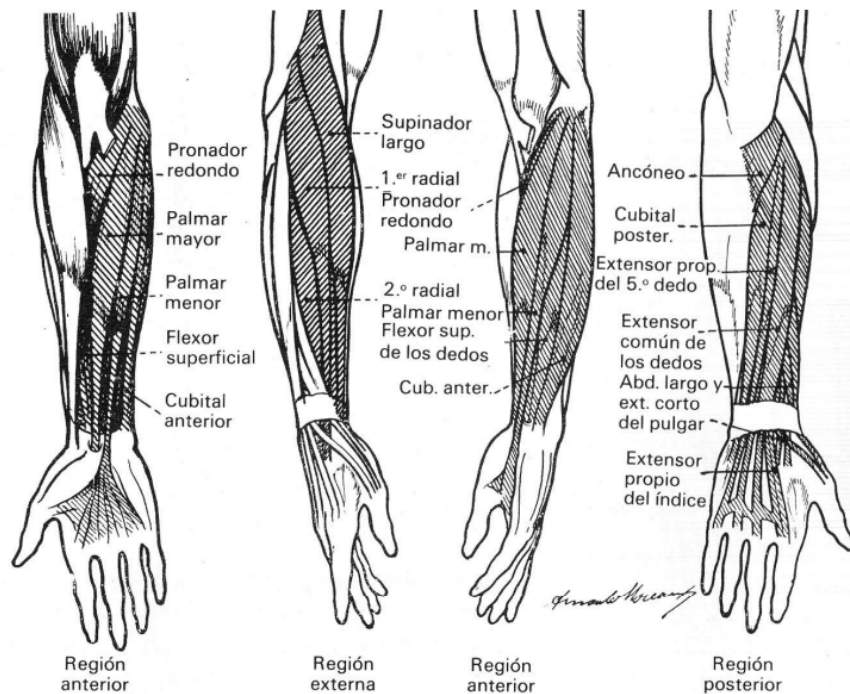


Figura 1.4. Composición muscular del antebrazo. Imagen tomada de [8].

Mano. La mano es un órgano que tiene una cara dorsal donde se encuentran los tendones de los músculos posteriores, una cara palmar que contiene los tendones de la región anterior y un conjunto de pequeños músculos agrupados en tres regiones que se describen a continuación:

1. Región tenar. Conformada por un grupo muscular encima del pulgar y está formada por el abductor corto, el oponente, el flexor corto y el abductor.
2. Región hipotenar. Grupo muscular interno del meñique y se compone del abductor, el flexor corto y el oponente, todos ellos del meñique.
3. Región palmar media. Son once músculos divididos entre interóseos y lumbricales. Los interóseos encuentran su origen entre los metacarpianos y se insertan en la primera falange de los dedos palmares. Los lumbricales son externos respecto a los interóseos. Se originan en los tendones del flexor común profundo y se insertan en la primera falange palmar.

A continuación, en la Figura 1.5, se presentan las diferentes capas musculares que conforman la mano.

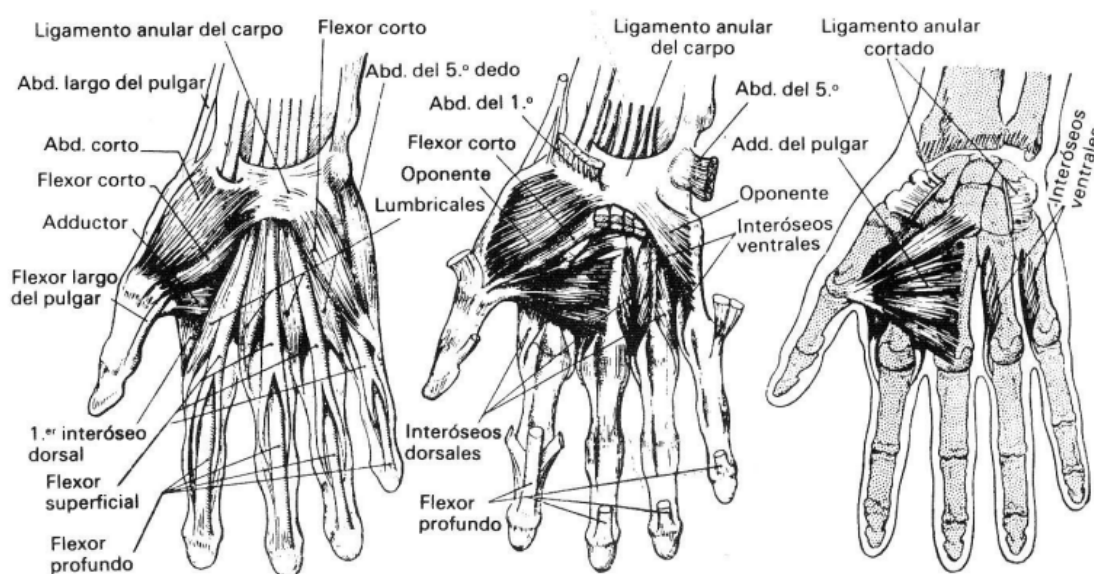


Figura 1.5. Composición muscular de la mano. Imagen tomada de [8].

1.2 Lesiones en miembros superiores

El estudio del M²S no solo es importante desde el punto de vista anatómico y fisiológico, también tiene aplicaciones vitales en el tratamiento y prevención de lesiones que pueden afectar significativamente la calidad de vida y capacidad funcional. A continuación, se proporciona una breve explicación de las lesiones más comunes que pueden ocurrir en los miembros superiores.

1.2.1 Lesiones de hombro

La *National Football League* (NFL) realizó un estudio a lo largo de 22 temporadas, desde 1980 hasta 2001 por medio del Sistema de Vigilancia de Lesiones de la NFL (NFLISS) para conocer las lesiones más comunes entre jugadores de Fútbol Americano. El estudio se realizó con personas en edad promedio de 25 años y se encontró que la lesión más frecuente del hombro en deportistas fueron la luxación acromioclavicular (41%), la inestabilidad del hombro glenohumeral anterior (20%) y las lesiones del manguito de los rotadores (12%) [10]. Un grupo de deportistas de alto riesgo son los que realizan lanzamientos supracefálicos los cuales suelen tener lesiones del labrum glenoideo y del manguito de los rotadores [7]. Entre estas lesiones, se describen aquellas que se generan durante los lanzamientos.

Luxación acromioclavicular. Es una lesión muy frecuente en deportistas que se presenta por una caída sobre el hombro con el brazo en aducción, lo que hará que se produzca una depresión del acromion. Cuando la clavícula no se fractura, la fuerza será absorbida por los ligamentos acromioclaviculares y los coracoclaviculares. Existen seis grados de luxación acromioclavicular:

- Tipo I: es un esguince de la articulación con los ligamentos acromioclaviculares y coracoclaviculares intactos.
- Tipo II: es un esguince moderado de la articulación con rotura de la cápsula articular (ligamentos acromioclaviculares) permaneciendo que los ligamentos coracoclaviculares estén intactos.
- Tipo III: la cápsula articular y los ligamentos coracoclaviculares están rotos. Existirá un mayor grado de desplazamiento de la clavícula distal.
- Tipo IV: los ligamentos coracoclaviculares están rotos con traslación posterior de la clavícula.
- Tipo V: ligamentos coracoclaviculares y fascia deltotrapezoidal rota con mayor desplazamiento que en el Tipo III.
- Tipo VI: disrupción total ligamentaria con desplazamiento inferior de la clavícula.

Inestabilidad del hombro glenohumeral anterior. La articulación glenohumeral (conocida como articulación del hombro) es una estructura que permite una amplia gama de movimientos en el hombro, pero requiere un equilibrio cuidadoso entre movilidad y estabilidad para funcionar correctamente y evitar lesiones. Por esta razón, es la articulación que sufre luxación con más frecuencia. Las lesiones más críticas causan inestabilidad en distintos músculos del hombro, lo que causa desgarros periósticos y aplasia glenohumeral. La inestabilidad anterior es relativamente frecuente en deportistas, siendo una secuela común tras una luxación anterior como la que se muestra en la Figura 1.6.

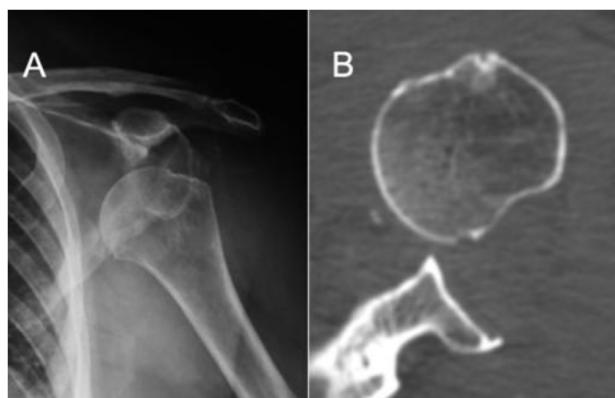


Figura 1.6. Imágenes de radiografía simple AP (A) y axial de TC (B) en paciente con luxación anterior de hombro. Se aprecia el desplazamiento anterior de la cabeza humeral respecto a la cavidad glenoidea. Imagen tomada de [11].

Otra lesión que suele asociarse a la luxación de hombro es la lesión de Bankart, que consiste en la rotura o desinserción del labrum anteroinferior, con desgarro asociado del periostio. En caso de que el labrum esté suelto, pero no desplazado, se denomina lesión de Perthes. Existe otra variante que es la lesión de Bankart óseo en la que se produce fractura del reborde óseo glenoideo anteroinferior.

Una lesión ósea muy prevalente, consecuencia de la luxación anterior, es la lesión de Hill-Sachs, que se muestra en la Figura 1.7. Se trata de una irregularidad ósea o muesca en la porción posterolateral de la cabeza humeral, como consecuencia del impacto de la misma con el reborde glenoideo anterior tras una luxación anterior.



I. MARCO TEÓRICO DEL M²S

Figura 1.7. Imagen coronal de artro-TC en paciente de 16 años con traumatismo deportivo y luxación anterior de hombro. Se aprecia lesión de Hill-Sachs (flecha) como consecuencia de la luxación. Existe también lesión del labrum superior (SLAP) (punta de flecha). Imagen tomada de [11].

Lesiones del manguito de los rotadores. Son lesiones generadas en el hombro y se refiere a un desgaste en el conjunto muscular que va de la escápula al húmero. Son más frecuentes en deportes de lanzamiento por encima de la cabeza (básquetbol, béisbol, lanzamiento de jabalina, entre otras) y en microtraumatismo repetidos.

Las lesiones del manguito de los rotadores son un proceso continuo. A medida que la lesión progrese en atletas de 25 a 40 años, se encontrará tendinopatía con fibrosis. En deportistas mayores de 40 años habrá roturas parciales y completas de dichos tendones. Son menos frecuentes los traumatismos directos del manguito rotador en el choque con otro deportista o las caídas con el brazo en extensión y abducción que pueden provocar un choque del supraespinoso bajo el acromion. En este caso pueden producirse roturas agudas del manguito de los rotadores sin pasar por el continuo: bursitis subacromio subdeltoidea, tendinopatía y rotura.

1.2.2 Lesiones del codo

Las lesiones del codo son comunes en los deportes de contacto como el Fútbol Americano debido a movimientos repetitivos que requieren los jugadores, especialmente aquellos en posiciones de lanzamiento como los quarterbacks. Estas lesiones pueden variar en gravedad desde esguinces leves hasta rupturas de ligamentos y fracturas. Son un área crítica de estudio y tratamiento debido a la alta incidencia y el impacto significativo en el rendimiento de los atletas. Una comprensión detallada de los tipos de lesiones, factores de riesgo y tratamientos disponibles es esencial para el manejo efectivo y la prevención de estas lesiones.

Lesiones ligamentosas. Las lesiones en el codo son las más frecuentes en lanzadores ya que están sometidos a diversas tensiones repetitivas en los ligamentos. Cada lanzamiento implica una alta velocidad y fuerza que impacta los ligamentos y tendones del codo, particularmente el ligamento colateral cubital (UCL). Esta repetición puede llevar a lesiones por sobreuso como

la tendinitis y las rupturas de ligamentos. Entre estas, una muy común es la lesión del ligamento colateral cubital, que es causada por el sobreuso en los deportes de lanzamiento por encima de la cabeza.

En deportes con mayor contacto, este estrés en valgo puede provocar lesiones agudas. De los tres haces del ligamento colateral cubital, el más importante estabilizador en valgo es el anterior. La rotura parcial o completa del haz anterior se conoce como codo del pitcher de beisbol. El haz transversal no tiene función estabilizadora debido a su origen e inserción en el cúbito. Las lesiones del ligamento colateral cubital se suelen acompañar de afectación del tendón común flexor. En casos de rotura completa, se referirá dolor con o sin chasquido durante el lanzamiento.

Luxación de codo. El 75% de luxaciones de codo ocurren durante la actividad deportiva. Los deportes de mayor riesgo en mujeres son la gimnasia y en hombres deportes de contacto, en especial el Fútbol Americano. La forma de clasificarlas es según el desplazamiento de cúbito y radio con respecto al húmero. Existen tres tipos principales:

- De cúbito y radio de forma conjunta desplazados anterior, posterior, lateral o medialmente (Figura 1.8).
- De cúbito anterior o posterior.
- De radio anterior, posterior o lateralmente.

El 90% de las luxaciones de codo son posteriores o posterolaterales. Este tipo de luxaciones no se asocian a ningún deporte en concreto ya que se producen en cualquier actividad que conduzca a la caída con la mano estirada. La fractura más habitual en adultos es de la cabeza radial, y tiene lugar en aproximadamente el 10% de los pacientes tras una luxación posterior. También se puede fracturar por luxación posterior. La combinación de luxación, fractura de cabeza de radio y fractura de la apófisis coronoides se conoce como triada terrible de codo. En casi todos los casos de luxación posterior completa, el ligamento colateral lateral cubital se ve afectado, soliendo romperse todos sus haces.



Figura 1.8. Reconstrucción 3D de TC con luxación medial cubitorradial. Imagen tomada de [11].

Lesiones osteocondrales y cuerpos libres intraarticulares. La osteocondrosis es un trastorno que afecta al centro de osificación de un niño o adolescente, ocasionando una falta de osificación endocondral. Este centro de osificación sufrirá un proceso de necrosis avascular con posterior reparación o secuelas. En el caso del codo, las localizaciones más frecuentes son la cabeza radial y el cóndilo humeral debido a fuerzas de compresión y cizallamiento que pueden producirse en jóvenes lanzadores o gimnastas (Figura 1.9).



Figura 1.9. Imágenes de radiografía simple AP (A) y RM coronal potenciada en T2 con supresión grasa (B) que muestra lesión osteocondral estable localizada en cóndilo humeral en joven deportista (flechas). Imagen tomada de [11].

Codo de golfista. El codo de golfista se refiere a una lesión por estrés del tendón flexor común y del origen del tendón pronador en el epicóndilo medial o epitroclea. Esta lesión también puede estar relacionada una técnica de golf inadecuada (golpear el suelo en lugar de la pelota) con la consiguiente desaceleración brusca.

Codo de tenista. El codo de tenista (epicondilitis) es una inflamación causada por el uso excesivo de los músculos extensores de la mano y la muñeca (como en el caso de golpear el revés de tenis, en especial a una mano). Afecta las fibras tendinosas del músculo extensor radial corto del carpo. Se producirá tendinosis, peritendinitis (edema del tejido adyacente) y desgarros parciales.

Lesión del tendón del tríceps. Las roturas del tendón del tríceps son raras en la práctica deportiva. Es más frecuente la existencia de un arrancamiento óseo en su inserción olecraneana, en especial en pacientes jóvenes (Figura 1.10).

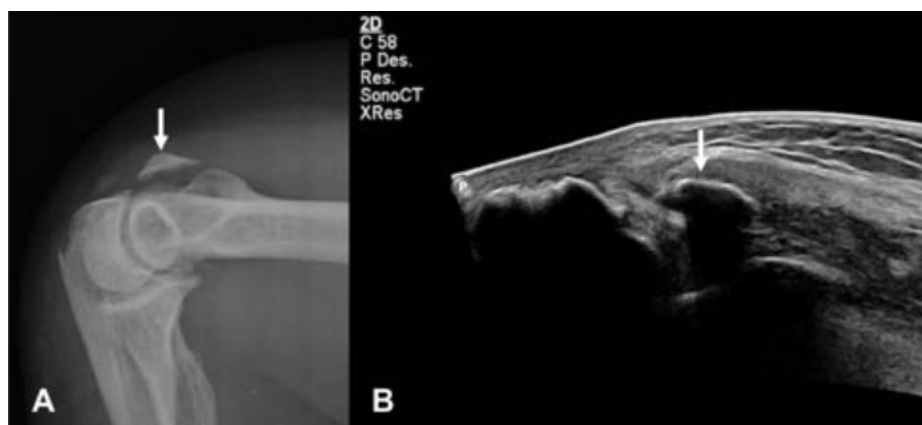


Figura 1.10. Imágenes de radiografía lateral de codo (A) y ecografía (B) con arrancamiento del olécranon (flecha). Imagen tomada de [11].

Codo de las pequeñas ligas. Es una fractura por avulsión del centro de osificación del epicóndilo medial producido por la tracción del tendón común flexor durante lanzamiento con técnica defectuosa con estrés en valgo excesivo. Una de sus causas deportivas más frecuentes es en lanzamientos en béisbol (Figura 1.11).

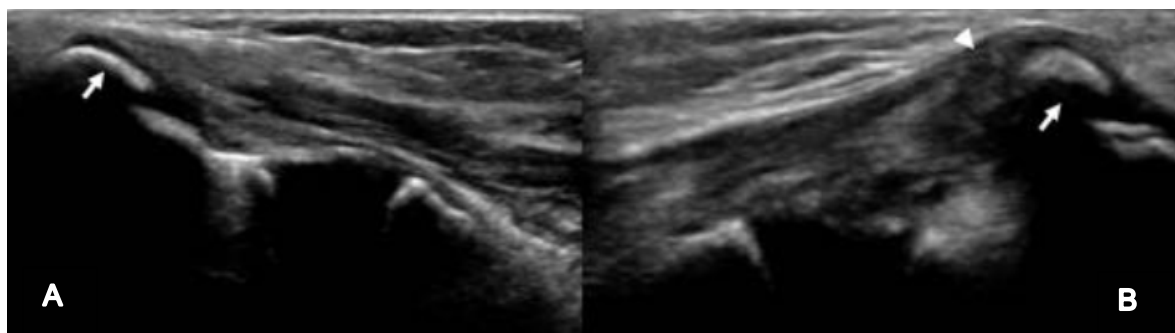


Figura 1.11. Imágenes ecográficas longitudinales de la epitróclea (flecha) y tendón común flexor en niño tras traumatismo deportivo, en codo asintomático (A) y sintomático (B). Se aprecia el despegamiento de la epitróclea en B y la alteración de la ecogenicidad del tendón común flexor (punta de flecha). Imagen tomada de [11].

Codo del pitcher de béisbol. La lesión más común en los pitcher es una rotura parcial o total de la banda anterior del ligamento colateral cubital del codo. Se produce por el estrés en valgo repetido durante las fases inicial y final del lanzamiento. A menudo se acompaña de otras lesiones en el codo, incluyendo contusión de la cabeza radial y cóndilo humeral, lesiones condrales de olécranon y neuropatía del nervio cubital.

Codo del portero de fútbol soccer. La lesión del portero se produce como resultado de los traumatismos repetitivos con hiper extensión del codo al bloquear la pelota. El olécranon chocará con la fosa olecraneana del húmero, produciendo daño en el cartílago, formación de osteofitos, y cuerpos libres intraarticulares.

1.2.3 Lesiones de muñeca y mano

Las lesiones de muñeca y mano se presentan con relativa frecuencia en deportes de contacto. La mayoría son contusiones y esguinces que se producen como consecuencia de una caída con la mano en extensión.

Fracturas. Las fracturas del radio distal y de los escafoides son frecuentes y suceden mediante el mecanismo de caída con la mano en extensión. Las fracturas del escafoides constituyen más del 60% de las lesiones del carpo y el principal síntoma es dolor en la tabaquera anatómica. La inflamación y el hematoma suelen estar ausentes. Una complicación común de las fracturas de escafoides es la necrosis avascular en algunas partes de la mano.

Articulación radiocubital distal. El principal estabilizador intrínseco de esta articulación es el complejo del fibrocartilago triangular (CFCT). El CFCT se compone de un disco central o menisco que se fija en la cara distal e interna de la epífisis radial y en la fosa y estiloides cubital. Otros componentes del complejo son los ligamentos radio cubital dorsal y palmar, el ligamento luno piramidal y su menisco. Los estabilizadores extrínsecos son el músculo pronador cuadrado, la membrana interósea y el extensor cubital del carpo. Las lesiones crónicas del CFCT pueden desarrollarse con el tiempo debido a movimientos repetitivos de la muñeca, específicamente la prona-supinación, el cual es el movimiento de rotación de la muñeca y el antebrazo. La pronación es cuando la palma de la mano gira hacia abajo y la supinación es cuando la palma gira hacia arriba. Movimientos repetitivos y continuos de prona-supinación pueden causar desgaste y desgarro en el CFCT.

Lesiones ligamentosas y tendinosas. Los principales ligamentos intrínsecos de la muñeca son el escafose milunar y el luno piramidal. Las lesiones de estos ligamentos en deportistas pueden pasar desapercibidas en un principio. Como efecto secundario a estas lesiones, puede existir una mecánica alterada de la muñeca por el dolor, lo que va degenerando de forma muy rápida la propia muñeca. La lesión de estos ligamentos es resultado de un evento traumático y las radiografías no suelen aportar hallazgos para el diagnóstico. La lesión del ligamento escafosemilunar se presenta con dolor, hinchazón y sensibilidad en el ligamento dorsorradial de la transición radiocarpiana. Las roturas pueden ser parciales o completas.

El ligamento interóseo lunopiramidal es más pequeño que el ligamento escafosemilunar. La rotura puede estar en relación con inestabilidad perilunar, choque ulnocarpal o fuerzas de carga en extensión máxima con desviación radial causada por muchos ejercicios de pronación en deportistas.

Las lesiones tendinosas de muñeca del deportista afectan con más frecuencia al extensor cubital del carpo, el primer compartimiento extensor (abductor largo y extensor corto del pulgar) y los tendones flexores. Las roturas podrán producirse de forma aguda o sobre un tendón ya dañado por tendinosis.

Pulgar del guardabosques o del esquiador. Recibe este nombre la lesión del ligamento colateral cubital del pulgar en la articulación metacarpo falángica como resultado de una fuerza radial con gran abducción del pulgar (por ejemplo, una caída donde el pulgar quede atrapado, como sucede con el fijador del bastón de esquí). En esta lesión, el ligamento colateral cubital puede deslizarse bajo el aductor del pulgar (lesión de *Stener*) o permanecer alineado con la cápsula articular (lesión no *Stener*). Normalmente, una lesión de *Stener* requerirá cirugía para situar el ligamento en posición anatómica y prevenir la inestabilidad articular.

Muñeca del remero. Se produce por la flexión y extensión repetida de la muñeca que causa una fricción entre los tendones del primer y segundo compartimento extensor de la muñeca. Esta entidad se conoce también como síndrome de la intersección. Los pacientes presentarán dolor y crepitación del antebrazo distal debido al líquido en las vainas tendinosas.

Muñeca del ciclista. Consiste en una neuropatía cubital causada por la compresión del nervio cubital en la mano y la muñeca como resultado de la presión directa sobre el nervio con el cuerno del manillar. Ocasionalmente, el manillar puede comprimir también el nervio mediano. En los ciclistas se le conoce como parálisis del manillar.

1.3 Rendimiento deportivo de miembros superiores

Las lesiones en los miembros superiores son comunes en muchos deportes y pueden afectar significativamente el rendimiento y la carrera de los atletas. A continuación, se presentan algunas estadísticas relevantes obtenidas de [12] [13] [14]:

Béisbol: los lanzadores son particularmente vulnerables a lesiones en el hombro y el codo. Aproximadamente el 30-40% de los lanzadores profesionales experimentan lesiones en el codo, como la lesión del ligamento colateral cubital (UCL por sus siglas en inglés), también conocida como la "lesión de Tommy John".

Tenis: las lesiones de codo, como la epicondilitis lateral (codo de tenista), representan alrededor del 20% de todas las lesiones en los tenistas. También, son comunes las lesiones en el hombro.

I. MARCO TEÓRICO DEL M²S

Natación: los nadadores de competición sufren frecuentemente de "hombro de nadador", una tendinitis del manguito rotador, con una prevalencia del 40-60% entre los nadadores de élite.

Baloncesto: las lesiones de hombro, como las luxaciones y las subluxaciones son comunes, con una incidencia que varía entre el 10-15% de todas las lesiones en el baloncesto.

Fútbol americano: las lesiones de hombro representan aproximadamente el 8-10% de todas las lesiones en la NFL. Las dislocaciones de codo son menos comunes pero significativas, representando alrededor del 1-2% de todas las lesiones. En la muñeca y la mano representan aproximadamente el 5-7% de todas las lesiones en la NFL.

Algunos de los factores de riesgo que causan lesiones son el uso de técnicas inadecuadas durante la práctica deportiva y el insuficiente fortalecimiento o flexibilidad de los músculos y tendones de los miembros superiores [15].

Estas estadísticas y análisis proporcionan una visión integral de la prevalencia y el impacto de las lesiones en los miembros superiores en los deportes. De esta manera, es posible conocer la magnitud del problema y así se pueden desarrollar mejores estrategias de prevención y tratamiento.

II. ANTECEDENTES

La literatura sobre el análisis de movimiento de miembros superiores es escasa en comparación con literatura del movimiento de otras partes del cuerpo humano. Por tal razón, los antecedentes de esta tesis se enfocan a revisar métodos relacionados al movimiento de varias partes del cuerpo humano y los antecedentes que proporcionaron información relevante en los experimentos de esta tesis. De acuerdo con dichos antecedentes, los métodos y trabajos relevantes en esta tesis se relacionaron a áreas tales como flujo óptico, inteligencia artificial y redes neuronales artificiales.

Este capítulo se divide de la siguiente manera. La sección 2.1 describe el flujo óptico. La sección 2.2 presenta la inteligencia artificial con enfoque en aprendizaje de máquina y aprendizaje profundo. Posteriormente, la sección 2.3 describe conceptos básicos de redes neuronales artificiales. Finalmente, la sección 2.4 reporta investigaciones y estudios recientes sobre las técnicas de medición de movimiento humano y de los miembros superiores.

2.1 Flujo óptico

El flujo óptico es un método de análisis de videos que permite determinar el movimiento de un objeto dentro de una secuencia de imágenes contenidas en un video (cuadro o fotograma) utilizando sus vectores de movimiento. Los vectores de movimiento son tensores utilizados para predecir el desplazamiento de las coordenadas de un bloque de un cuadro y las coordenadas de un bloque de un cuadro anterior. Según Al Bovik, el flujo óptico, es un método que calcula el movimiento de las intensidades de una imagen, que puede ser atribuido al movimiento de objetos en una escena en un video. Aunado a esto, los métodos de flujo óptico se basan en cálculos del movimiento de las intensidades de la imagen en un tiempo determinado de un video [16].

Qihong Ke et al. analizaron los diferentes métodos de flujo óptico en [17] y concluyen que existen muchas formas de estimar el flujo óptico entre dos cuadros (*frames*), siendo el método diferencial (cálculo de vectores) el más usado. Este método asume que la intensidad de la

II. ANTECEDENTES

imagen es constante entre los *frames* de un video. Esto se puede definir de la siguiente manera: dado un video, la intensidad de un pixel volumétrico (voxel) en position (x, y) del t -ésimo *frame*, $I(x, y, t)$, y dada la hipótesis de que la intensidad de la imagen es constante, entonces, la intensidad del voxel también será la misma a pesar de pequeños cambios de posición o periodos de tiempo. Esta hipótesis se puede plantear de la siguiente manera:

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t) \quad (2.1)$$

Donde $(\delta x, \delta y, \delta t)$ es el cambio de movimiento. Esta expresión se puede representar con una serie de Taylor de la siguiente manera:

$$I(x + \delta x, y + \delta y, t + \delta t) \approx I(x, y, t) + \frac{\partial I}{\partial x} \delta x + \frac{\partial I}{\partial y} \delta y + \frac{\partial I}{\partial t} \delta t \quad (2.2)$$

Por lo tanto:

$$\frac{\partial I}{\partial x} \delta x + \frac{\partial I}{\partial y} \delta y + \frac{\partial I}{\partial t} \delta t \approx 0, \quad (2.3)$$

$$\frac{\partial I}{\partial x} U_x + \frac{\partial I}{\partial y} U_y + \frac{\partial I}{\partial t} \approx 0 \quad (2.4)$$

Donde $U_x = \frac{\delta x}{\delta t}$, $U_y = \frac{\delta y}{\delta t}$ son dos componentes del flujo óptico del voxel, o pixel, (x, y, t) .

Existen dos variantes principales al estimar el flujo óptico de un video: flujo óptico disperso y flujo óptico denso.

2.1.1 Flujo óptico disperso

El flujo óptico disperso selecciona un conjunto de pixeles dispersos para rastrear sus vectores de movimiento. Dichos pixeles tienen información de características de bordes y esquinas que se pasan a la función de cálculo de flujo óptico de *frame a frame* para garantizar que se sigan los mismos puntos. Un ejemplo de implementación de cálculo de flujo óptico disperso es el algoritmo de *Lucas-Kanade*. Este algoritmo consiste en resolver la ecuación del flujo óptico

II. ANTECEDENTES

(ecuaciones 3 y 4) para todos los píxeles de esa zona de 3x3 alrededor de píxeles dispersos, usando el criterio de los cuadrados mínimos [18] [19].

2.1.2 Flujo óptico denso

El flujo óptico denso calcula un vector de movimiento para cada uno de los píxeles de los *frames* del video. Aunque el cálculo puede ser más lento que el flujo óptico disperso, este proporciona un resultado más preciso, ideal para aplicaciones de segmentación de videos y clasificación de movimientos en video.

Un ejemplo de flujo óptico disperso es el algoritmo de *Gunnar-Farneback*. Este algoritmo se puede resumir de la siguiente manera: primero se aproximan bloques de píxeles de dos *frames* utilizando polinomios cuadráticos definidos por una transformación de expansión polinomial [20]. Segundo, observando cómo los polinomios se transforman bajo movimiento, se utiliza un método para estimar el desplazamiento utilizando los coeficientes de dicha expansión polinomial. Finalmente, se obtiene el flujo óptico denso después de una serie de arreglos y refinamientos [21].

2.2 Inteligencia artificial

La inteligencia artificial (IA) nació a mediados del siglo XX y es una disciplina que persigue la construcción de máquinas y programas capaces de realizar tareas complejas con una habilidad y eficiencia iguales o superiores a las de un ser humano [22]. A continuación, se presentan los conceptos más relevantes de IA.

2.2.1 Ciencia de datos

La ciencia de datos se refiere a las técnicas necesarias para comprender y manejar los conjuntos de datos. Está diseñada para descubrir correlaciones en datos que representen fenómenos reales o abstractos [23]. El proceso de la ciencia de datos conlleva los siguientes pasos:

Recolección de datos. Adquirir datos por medio de sensores o fuentes como bases de datos, sensores, servidores, archivos de texto, imágenes, videos, biopotenciales, audio, etc.

II. ANTECEDENTES

Preprocesamiento de datos. Limpieza de datos que no son útiles, eliminar ruido y transformarlos para posterior tratamiento.

Análisis exploratorio de datos (EDA). Analizar la composición de los datos para conocer las propiedades y características contenidas en ellos.

Modelado y algoritmos. Uso de modelos estadísticos y de aprendizaje automático para constituir modelos de regresión o predictivos. En esta etapa es común el uso de lenguajes de programación como Python, SQL, R y herramientas como TensorFlow o Pythorch.

Implementación. Desarrollo de dispositivos y sistemas que puedan analizar los datos de forma inteligente.

Los modelos y métodos de la ciencia de datos evolucionan rápidamente a lo largo de su diseño, desarrollo e implementación por medio de aprendizaje dinámico. Los modelos y métodos se entrenan utilizando aprendizaje semiautomático, aprendizaje por refuerzo y el meta-aprendizaje [24].

2.2.2 Aprendizaje de máquina

El Aprendizaje de máquina o automático (ML por sus siglas en inglés) es una forma de inteligencia artificial que aprende a partir de los datos. De forma concreta, esta disciplina trata de crear modelos y algoritmos capaces de generalizar comportamientos y reconocer patrones a partir de una información suministrada con ejemplos [25]. Las formas de aprendizaje más básica del ML son las siguientes:

Aprendizaje supervisado (*Supervised Learning*). El aprendizaje supervisado se basa en entrenar al modelo con un conjunto de datos que se componen de muestras con salidas conocidas. Con este conjunto, el modelo ajusta sus parámetros para reducir la magnitud de una función de costo. Es necesario permitir que el modelo desarrolle una capacidad de generalización y evitar un problema llamado sobreajuste, el cual evita que el modelo logre predecir correctamente las muestras fuera del conjunto de muestras de entrenamiento [26]. Dentro de la categoría de aprendizaje supervisado, existen dos formas de entrenar: regresión y

II. ANTECEDENTES

clasificación. El objetivo de la regresión es predecir un resultado que varía en un rango continuo a partir de la información proporcionada por una serie de variables predictoras. Por otra parte, la clasificación se centra en predecir resultados en categorías discretas [25].

Aprendizaje no supervisado (*Unsupervised Learning*). El aprendizaje no supervisado utiliza un conjunto de datos que se componen solo de las muestras de entrada para encontrar la organización de estos y separarlos en grupos. Una de las técnicas más usadas de aprendizaje no supervisado es el Agrupamiento (*Clustering*), donde el algoritmo agrupa las muestras de entrada según sus características estadísticas. Otra técnica que se puede utilizar es Detección de Anomalías (*Anomaly Detection*), donde el algoritmo aprende a partir de los datos de entrenamiento cuáles son los valores normales. A partir de estos valores, el algoritmo indica si algún elemento cae fuera de los rangos normales [25].

2.2.3 Aprendizaje profundo

El aprendizaje profundo (DL por sus siglas en inglés) es un subcampo del ML que se basa en el diseño de modelos multicapa con una estructura jerárquica organizada de una forma similar a la estructura neuronal del cerebro humano. Los modelos se entrenan utilizando conjuntos masivos de datos y son generalmente arquitecturas de redes neuronales artificiales que contienen muchas capas para conocer las propiedades abstractas de los datos que se van aprendiendo [25].

El aprendizaje profundo permite que la computadora construya conceptos complejos a partir de conceptos más simples. Hay dos formas principales de medir la profundidad de un modelo. La primera se basa en el número de instrucciones secuenciales que se deben ejecutar para evaluar la arquitectura. Otro enfoque, considera la profundidad de un modelo como la profundidad del grafo que describe cómo se relacionan los conceptos entre sí [27].

2.3 Redes neuronales artificiales

Las redes neuronales artificiales (ANN, por sus siglas en inglés) son los modelos más utilizados en ML y DL y se componen de unidades de procesamiento simples inspiradas en el

II. ANTECEDENTES

comportamiento y estructura de las neuronas biológicas. Las ANN imitan los mecanismos del cerebro y el sistema nervioso, de manera que realizan lo siguiente:

1. El conocimiento es adquirido por la red mediante un proceso de aprendizaje.
2. Las fuerzas de conexión interneuronal, conocidas como pesos sinápticos son usados para almacenar el conocimiento adquirido.
3. El algoritmo de aprendizaje es el procedimiento para modificar los pesos sinápticos de la red de manera ordenada y así lograr un objetivo de diseño deseado [28].

Existen varias arquitecturas de redes neuronales, y entre las más populares en la ciencia y la industria se encuentran las redes neuronales convolucionales (CNN por sus siglas en ingles). La arquitectura de las CNN está conformada por capas convolucionales, max-pooling, capas densas y función de activación. A continuación, se describen cada una de las capas de una CNN.

Capas de convolución. La convolución es una operación de suma de productos que genera mapas de características. Es un proceso que consiste en implementar una operación lineal e invariante en el tiempo entre un conjunto de kernels o filtros con la entrada de la red o la salida de una capa anterior [29]. Para una entrada I y un filtro (o kernel) K se tiene:

$$f[(I * K)(m, n)] = f \left[\sum_m \sum_n I(i - m, j - n) \cdot K(m, n) \right] \quad (2.5)$$

Donde $I(m, n)$ es la imagen de entrada o de la capa anterior, $K(m, n)$ es el valor de la posición (m, n) del kernel, m y n son los índices que recorren las dimensiones del kernel y $f[u]$ es la función de activación.

Capa max-pooling o capa de reducción. El agrupamiento extrae estadísticas como el máximo de una ventana fija del mapa de características. Esto reduce el número de elementos para representar las características, reduce el problema del sobreajuste y reduce la carga computacional. La operación está dada por:

$$O(i, j) = \max \{f(m, n) | (m, n) \in W_{ij}\} \quad (2.6)$$

II. ANTECEDENTES

Donde $O(i, j)$ es el valor en la posición (i, j) de la salida después de la operación de *max-pooling*, W_{ij} es la ventana de *pooling* centrada en la posición (i, j) y $f(m, n)$ son los valores de la salida de la capa convolucional dentro de la ventana W_{ij} .

Capas densas o capas de clasificación (*fully connected*). Estas capas calculan la pertenencia de una muestra a cada clase mediante las características extraídas de las capas de convolución y *max-pooling*. Para ello, esta capa relaciona todas las características que se han extraído previamente al entrenamiento de la red neuronal para obtener una clasificación final. Para una capa densa en una red neuronal, la operación se puede describir mediante la siguiente ecuación:

$$y = f(W(m, n)I(m, n) + b) \quad (2.7)$$

Donde $I(m, n)$ la capa anterior, $W(m, n)$ es el peso de la posición (m, n) del kernel, b es el *bias* de la neurona y $f(u)$ es la función de activación.

Función de activación. Las funciones de activación se encuentran presentes en las capas de convolución y en las capas *fully connected* y se utilizan para varios propósitos tales como generar características con modelado no lineal o mejorar el aprendizaje. Existen varios tipos de funciones de activación, pero las más populares son la sigmoideal, *softmax* y ReLu.

La función sigmoideal se emplea en la clasificación binaria, donde 0 es una respuesta negativa y 1 es una respuesta positiva. La función sigmoideal está definida por la ecuación:

$$\sigma(u) = \frac{1}{1 + e^{-u}} \quad (2.8)$$

Donde $\sigma(u)$ es la salida de la función sigmoideal y u es la entrada a la neurona (también conocido como el valor activado previo).

En el caso de tener una clasificación multiclase, se utiliza la función *softmax* con entropía cruzada, ya que permite obtener una probabilidad normalizada y es una extensión de la función sigmoideal. Para un vector de entrada $u = [u_1, u_2, \dots, u_K]$ la función *softmax* se define como:

$$\sigma(u)_i = \frac{e^{u_i}}{\sum_{j=1}^K e^{u_j}} \quad (2.9)$$

Donde $\sigma(u)_i$ es la i -ésima componente del vector de salida después de aplicar la función *softmax*, u_i es la i -ésima componente del vector de entrada u y K es el número de clases (o la dimensión del vector z).

La función ReLU (Función de Rectificación Lineal) se utiliza para generar redes profundas, ya que acelera el entrenamiento en un factor de 6 comparado con las funciones sigmoide y evita el problema del desvanecimiento de gradiente cuando se presentan muchas capas [30]. Esta función está dada por:

$$f(u) = \max(0, u) \quad (2.10)$$

Donde $f(u)$ es la salida de la función ReLU y u es la entrada a la neurona.

2.4 Análisis del estado del arte

Se ha prestado mucha atención al seguimiento del movimiento corporal desde la década de 1990, ya que el rastreo del movimiento humano genera conocimientos para diseñar dispositivos y algoritmos que describen la mecánica del cuerpo. El interés en este tema parte de sus diversas aplicaciones en ingeniería biomédica, entrenamiento deportivo, realidad virtual, cooperación humano-robot, etc. Por ello, existen diversas técnicas para medir el movimiento humano, incluyendo sistemas de seguimiento óptico, mecánico, magnético, acústico y magnético-inercial.

El análisis del estado del arte se divide en técnicas de movimiento humano, ciencia de datos aplicada al análisis del movimiento humano, el cual se delimita en: reconocimiento de humanos, reconocimiento de actividades, reconocimiento de deportes y reconocimientos de miembros superiores.

2.4.1 Técnicas de medición de movimiento humano

El estudio [31] tiene como objetivo mapear la práctica actual para la evaluación del miembro superior en sociedades profesionales internacionales y laboratorios afiliados para identificar barreras que puedan limitar el análisis del movimiento del miembro superior en la práctica clínica. Se utilizó una encuesta en línea para recopilar respuestas a través de coordinadores de sociedades internacionales de análisis del movimiento en la que el 75% de los laboratorios realizaron algún tipo de evaluación para miembros superiores. Los métodos de evaluación más utilizados fueron el análisis de movimiento en 3D, electromiografía de superficie (sEMG), examen clínico y observación visual, video en 2D y examen clínico apoyado por instrumentación. En los laboratorios se evaluaron 28 movimientos y 18 tareas funcionales. Además, se encontró que la práctica en estos estudios eran variables, sin enfoques identificados en cuanto a los criterios de grupos de pacientes vistos, los métodos de evaluación, los resultados, los movimientos, las tareas funcionales, los conjuntos de marcadores y los modelos utilizados. La conclusión discute la necesidad de estandarizar la práctica y desarrollar consensos en la evaluación del miembro superior para mejorar su uso en la práctica clínica, así como en la investigación.

2.4.2 Ciencia de datos para análisis de movimiento humano

En [32] se presentan dos conjuntos de datos que contienen mediciones de sensores de Magnetismo, Tasa angular y Gravedad (MARG). Estos datos permiten analizar la canalización completa de la pose inercial a la pose corporal completa. Se recolectaron varias pruebas en las que se realizaron seis tipos de secuencias diferentes, incluyendo actividades cotidianas y movimientos aleatorios. Según los experimentos, los datos son muy variables y abarcan casi todo el rango de movimiento humano, a la vez que incluyen fuentes comunes de error encontradas en condiciones reales. Según los resultados, se generó una base de datos con conjuntos de datos con 3.5 millones de muestras y se presentó un método de pose inercial utilizado para validar la calidad de los datos en ambas adquisiciones.

En [3], Zhang et al. midieron la eficiencia de los sensores magnético-inerciales (*Inertial Measurement Unit and Magnetic sensor*, MIMUs), que utilizan una combinación de giroscopio

II. ANTECEDENTES

de tres ejes que mide la velocidad angular, acelerómetro, la aceleración lineal y magnetómetro, el cual mide la intensidad del campo magnético. Al combinar estas mediciones, los sensores magnético-inerciales pueden estimar la orientación de los segmentos corporales humanos en tiempo real. Los autores presentaron un enfoque de filtrado Kalman adaptativo y prueba de hipótesis para reducir el ruido de las señales adquiridas. Los resultados del estudio muestran que el enfoque propuesto supera a los métodos existentes en términos de precisión y robustez en varios escenarios, incluyendo movimientos estáticos y dinámicos.

El artículo [33] presenta un estudio de un sistema de seguimiento basado en unidades de medición inercial de microelectromecánica (MEMS). El sistema consta de siete unidades de medición inercial MEMS y proporciona una solución para rastrear el movimiento articular. El estudio desarrolló un nuevo sistema de seguimiento angular para monitorear los ángulos articulares durante los ejercicios de rehabilitación y se probó en 10 sujetos humanos realizando ejercicios de caminar, trotar y hacer zancadas. El sistema de seguimiento se comparó con un sistema de seguimiento óptico disponible comercialmente. Los resultados mostraron que el sistema IMU era comparable al sistema de seguimiento óptico en términos de precisión y podría usarse potencialmente para la rehabilitación en exteriores o en el hogar.

En [34] se describe un estudio que analiza los datos electromiográficos (EMG) de los movimientos de brazos para distinguir entre las activaciones musculares fásicas y tónicas durante la aceleración, deceleración y estabilización. Las pruebas estadísticas utilizadas fueron pruebas ANOVA de un solo factor con las activaciones musculares fásicas y tónicas como factores, y el nivel de significación se estableció en 0.05. El estudio también acopló medidas de EMG con la cinemática para probar si había diferencias en los ángulos cinemáticos a lo largo de diferentes direcciones. Los resultados del análisis de datos se utilizaron para proporcionar una caracterización completa de las activaciones de EMG en movimientos de extremidades superiores punto a punto y para investigar las relaciones entre EMG fásicas y tónicas.

Un método para detectar la actividad muscular del brazo superior y clasificar los movimientos del antebrazo, se presenta en el artículo [35]. Dicho método utiliza sensores de resistencia sensible a la fuerza (FSR) colocados en el brazo superior. El estudio muestra que el

II. ANTECEDENTES

músculo del brazo superior puede proporcionar niveles claros de actividad para clasificar los movimientos del antebrazo y el método puede ahorrar esfuerzo y tiempo en la recopilación de datos de entrenamiento.

Una técnica que se valida en la industria [4] tiene sus bases en Unidades de Medida Inercial (IMUs), para estimar en tiempo real la cinemática de los miembros superiores durante tareas típicas de ensamblaje industrial. Se colocaron tres IMUs en la parte superior del cuerpo de seis participantes para medir los ángulos del hombro y del codo durante tareas de "tomar y colocar" en una fábrica. Los resultados se validaron mediante un sistema de captura de movimiento estereofotogramétrico. Los errores asociados con la técnica propuesta fueron moderados y confirmaron su idoneidad para un contexto de colaboración humano-robot en la industria.

Reconocimiento de humanos

El artículo [36] propone un método de captura de movimiento sin marcadores para entornos de diversas personas utilizando múltiples cámaras calibradas y sincronizadas. El método predice la pose 3D de cada persona. Esto se usa para predecir el cuadro delimitador de cada imagen de la cámara en el siguiente fotograma, lo que da un efecto sinérgico en el rendimiento de la captura de movimiento de vídeo. La reconstrucción sinérgica se basa en la combinación de la predicción y el filtrado espaciotemporal basado en el modelo esquelético humano, lo que permite la reconstrucción 3D precisa y suave de la persona. Este método se evaluó utilizando varios conjuntos de datos y un campo deportivo real, demostrando una alta precisión y suavidad. En particular, el método logró una precisión de 31.5 mm en MPJPE y 99.5% en PCP en un entorno con cinco personas en movimiento dinámico. Además, el método logró reconstruir con precisión los movimientos detallados de todos los jugadores en un partido de fútbol utilizando solo unas pocas cámaras.

Reconocimiento de actividades

La estandarización de diversos métodos para un solo enfoque se propone en [37], donde se combina la detección y el seguimiento para la estimación de la postura humana en videos. Se presentan tres componentes principales: la Red de seguimiento de clips, el *Pipeline* de

II. ANTECEDENTES

seguimiento de videos y el procedimiento de fusión espacio-temporal. La Red de seguimiento de clips realiza la detección y el seguimiento de las articulaciones del cuerpo simultáneamente en pequeños clips de video, mientras que el Pipeline de seguimiento de videos fusiona *tracklets* de longitud fija a *tracks* de longitud arbitraria. Finalmente, el procedimiento de fusión espacio-temporal perfecciona las ubicaciones de las articulaciones en función de la información espacial y temporal. Al combinar estas técnicas, se pueden manejar escenarios complejos con varias personas. El enfoque propuesto direcciona hacia la investigación futura en la estimación y seguimiento de la postura humana en videos.

AlphaPose presentado en [1], es un sistema para la estimación y seguimiento de poses de múltiples personas en tiempo real. El sistema utiliza varias técnicas, incluyendo una CNN que toma una imagen para generar un mapa de calor, un algoritmo para eliminar detecciones redundantes y un identificador de pose para la estimación conjunta de poses y seguimiento. También incluye un generador de propuestas guiado por partes e incorpora datos de diferentes dominios en el entrenamiento del modelo, para mejorar su capacidad para generalizar y adaptarse a diferentes situaciones.

Un método de estimación y seguimiento de la postura humana en videos utilizando una red neuronal de grafos (GNN) se describe en [5], el cual considera información espacial-temporal y visual. El método utiliza una técnica para modelar las relaciones entre las articulaciones en diferentes cuadros y predice las posturas en un cuadro a partir de un historial de seguimiento sin mirar ninguna pista de detección. El método se evalúa con las bases de datos PoseTrack 2017 y PoseTrack 2018. La visualización de las atenciones entre las articulaciones aprendidas por el modelo GNN muestra que las articulaciones en cuadros consecutivos influyen en la postura actual, las conexiones dentro del mismo cuadro modelan la dinámica espacial y las conexiones entre cuadros incorporan la dinámica temporal.

El artículo [38] se centra en el desarrollo de una red neuronal Wavelet (WNN) para mejorar para mejorar el rendimiento en la decodificación de movimientos a partir de señales EEG de MI. Se utilizó el conjunto de datos EEG público de BNCI Horizon 2020, que contiene sujetos de entre 22 y 40 años. Quince sujetos sanos participaron, realizando seis movimientos

II. ANTECEDENTES

funcionales de un solo miembro superior: pronación/supinación del antebrazo, apertura/cierre de la mano y flexión/extensión del codo. La metodología incluyó el preprocesamiento de señales, segmentación de datos, extracción de características, estandarización y selección, y clasificación basada en WNN. El estudio optimizó varios factores como la ondícula madre, la función de ondícula multidimensional, el número de canales y el tiempo del segmento de imagen a través de experimentos. La mejor precisión lograda fue del $86.27 \pm 6.98\%$ utilizando la ondícula madre *coif1* optimizada con una descomposición de seis niveles. También se exploró el impacto de reducir el número de canales en la precisión y la eficiencia computacional. El estudio tuvo como objetivo optimizar el método de experimentación y comparar los resultados para establecer la efectividad del enfoque WNN propuesto.

TA3DNet, una red neuronal para el reconocimiento de acciones humanas [39], es una red 3D de atención temporal débilmente supervisada diseñada para acelerar las redes neuronales convolucionales 3D (3D CNN). Utiliza una selección de marcos temporales para reducir la redundancia temporal en los videos, eligiendo cuadros a corto plazo con conexiones a largo plazo. Se aplica un módulo de atención temporal para asignar diferentes pesos a cada cuadro. Este módulo se entrena de manera débilmente supervisada, actualizando los pesos basándose únicamente en etiquetas de clase sin información de eventos ni etiquetas adicionales. TA3DNet demostró un alto rendimiento en dos conjuntos de datos, UCF101 y HMDB51. TA3DNet consta de 3 módulos W-STAM, 9 módulos de inicio, 4 capas convolucionales simples y 5 capas de agrupamiento. Las capas convolucionales toman la normalización por lotes y la unidad lineal rectificada (ReLU). Se utilizaron técnicas de aumento de datos y se entrenó la red en una estación NVIDIA DGX con cuatro GPUs Tesla V100. Se utilizaron diferentes estrategias de aprendizaje para cada conjunto de datos. Superó a los modelos existentes en cuanto a precisión, eficiencia y capacidad de procesamiento, aunque utilizó menos cuadros de entrada.

Una metodología innovadora para el reconocimiento de acciones humanas 3D utilizando mapas de profundidad se presenta en [40]. La propuesta se centra en un marco de trabajo eficiente y directo que considera las dependencias a corto y largo plazo en la representación de las acciones humanas. Se desarrolla una característica a nivel de cuadro, llamada vector de

II. ANTECEDENTES

gradiente orientado a la profundidad (DOGV) para capturar la apariencia y el movimiento a corto plazo. Para las dependencias a largo plazo, se construyen CNN para agregar características a nivel de cuadro en la secuencia espacial y temporal. El método se evalúa exhaustivamente en cuatro conjuntos de datos de referencia públicos, incluyendo NTU RGB+D, NTU RGB+D 120, PKU-MMD y UOW LSC. El estudio concluye que el enfoque propuesto es eficaz para reducir la redundancia temporal a nivel de cuadro y resaltar la información temporal principal. La capacidad del método para funcionar bien en una variedad de conjuntos de datos indica su robustez y generalización, lo que es crucial para aplicaciones prácticas en reconocimiento de acciones.

Para mejorar la eficiencia de la extracción de características de flujo óptico y explorar métodos de fusión de características espaciotemporales, el artículo [41] propone una red de reconocimiento de acciones humanas rápida. El modelo se entrena de forma integral, utilizando MotionNet para generar flujos ópticos y CNN con OFF (*Optical Flow Guided Feature*) para clasificación. Combina características espaciales y temporales en características de fusión. El modelo propuesto solo necesita entradas RGB, lo que simplifica el proceso y aumenta la velocidad para lograr una precisión en cada base de datos que van específicamente del: 91.5% en UCF-101 dataset, 67.9% en HMDB51 dataset, 83.3% en MSR Daily Activity3D dataset y 91.25% en Florence 3D action dataset. El modelo puede alcanzar 140 fotogramas por segundo usando CNN para calcular el flujo óptico, lo que representa una mejora significativa en velocidad comparado con métodos anteriores. Se entrena de manera integral, permitiendo que una sola red aprenda información tanto espacial como de movimiento.

Un tema que ha sido estudiado actualmente es el reconocimiento de acciones humanas en entornos de colaboración humano-robot (HRC) en la manufactura inteligente. El artículo [42] propone un modelo de red neuronal convolucional de dos flujos (CNN) transferible para mejorar la eficiencia operativa. El modelo se basa en el flujo óptico y el aprendizaje por transferencia para analizar simultáneamente la información espacial y temporal de las acciones humanas. Se utiliza un modelo de CNN preentrenado en un conjunto de datos de acciones humanas no específico de la manufactura y luego se transfiere para reconocer acciones en un

II. ANTECEDENTES

dominio objetivo con muestras de entrenamiento limitadas. La investigación se centra en superar desafíos como la dificultad de obtener grandes volúmenes de datos para el entrenamiento de redes neuronales en entornos de manufactura realistas. Se extrae información espacial y temporal de acciones humanas mediante un modelo de dos flujos que analiza simultáneamente ambos tipos de datos. Se evalúa la efectividad del modelo desarrollado utilizando el montaje de bloques de motor como escenario representativo de manufactura. El estudio concluye que la incorporación de información espacial y temporal mejora la precisión del reconocimiento de acciones humanas y el aprendizaje por transferencia es efectivo para adaptar la capacidad de extracción de características de la red de un dominio fuente a un dominio objetivo.

Una técnica que busca capturar información temporal en múltiples estados y escalas para un análisis efectivo del movimiento se presenta en [43], el cual es un método de reconocimiento de acciones humanas basado en el módulo de Excitación de Movimiento y Agregación Temporal (META). Las bases de datos utilizadas son: UCF101 y HMDB-51. Los videos provienen de fuentes como *YouTube* y *Google Video* y abarcan una amplia gama de movimientos como sonreír, reír, masticar, hablar, esgrima, abrazar, patear, besar y estrechar manos. El método propuesto utiliza muestreo relacional temporal en fotogramas de video y se compone de dos módulos principales: Módulo de Excitación de Movimiento a Múltiples Escalas (MME) que fortalece la relación temporal entre fotogramas de video y el Módulo de Agregación Temporal de *Squeeze* y Excitación (SETA) que transforma la convolución local en un conjunto de sub-convoluciones para extraer características a diferentes escalas temporales. Se realiza una combinación de características de flujo óptico y RGB para el reconocimiento de acciones. Se hace una fusión espaciotemporal y clasificación, con un preentrenamiento cruzado de modalidades para mejorar la utilización de características de flujo óptico. El método alcanza una precisión del 96.0% en UCF101 y 71.2% en HMDB-51, superando otros métodos contemporáneos.

El artículo [44] se centra en el reconocimiento de acciones humanas en escenarios de rescate de emergencia utilizando una red de fusión espaciotemporal. El trabajo aborda los desafíos de

II. ANTECEDENTES

la utilización insuficiente de información temporal y espacial en videos, lo que conduce a tasas de precisión más bajas en el reconocimiento de acciones. Una contribución clave del documento es el desarrollo de una red de fusión asíncrona de dos flujos basada en Redes de Segmentos Temporales (TSN) y Memoria a Corto y Largo Plazo Bidireccional (Bi-LSTM) para la clasificación de secuencias de video completas en situaciones de rescate de emergencia. La metodología involucra una estrategia de segmentación de dominio temporal basada en muestreo aleatorio para mantener la estructura general del dominio temporal del video. Este enfoque considera la relación asíncrona espaciotemporal e incluye múltiples secuencias de movimiento asíncronas como entrada para la red de convolución temporal. Las características espaciotemporales se fusionan a nivel de la capa de convolución para reducir la pérdida de características. Luego, las características de fusión espaciotemporal de capa media se procesan a través de Bi-LSTM para obtener características de movimiento humano a lo largo de la dimensión temporal del video. Los experimentos realizados con datos mejorados de rescate de emergencia y datos UCF101 demuestran que el método propuesto fusiona completamente la información espacial y temporal, mejorando significativamente la precisión del reconocimiento de acciones humanas en escenas dinámicas. Además, el método propuesto es más rápido que los métodos tradicionales.

En [45] se presenta un enfoque de detección de caídas que explora la geometría corporal humana disponible en diferentes fotogramas de la secuencia de video, utilizando la estimación de la postura, el ángulo y distancia del vector formado por el centroide de la cabeza de la imagen facial identificada y el centro de la cadera del cuerpo, y el vector alineado con el eje horizontal del centro de la cadera. Utilizan un clasificador de Máquinas de Soporte Vectorial (SVM) de dos clases y una Red de Convolución Temporal (TCN) entrenadas en las nuevas imágenes de características construidas. Al mismo tiempo, se entrena una red de Memoria a Corto y Largo Plazo (LSTM) en las secuencias calculadas de ángulos y distancias para clasificar actividades de caída y no caída. Se realizan experimentos en el conjunto de datos FD de Le2i, que incluye 221 videos, con 131 de caídas y 90 de actividades diarias de la vida (ADLs), y el conjunto de datos FD de UR, Contiene 70 secuencias (30 caídas + 40 ADLs), grabadas con dos cámaras Kinect desde diferentes perspectivas, resultando en 60 secuencias de caídas y 40 de actividades

II. ANTECEDENTES

no caídas. La evaluación cruzada con diferentes conjuntos de datos resalta la robustez y la aplicabilidad general del método.

Reconocimiento de deportes

Encaminado hacia los deportes, el artículo [46] ofrece una revisión exhaustiva del Reconocimiento de Acciones Humanas (HAR) utilizando la Visión por Computadora (CV). Se aprovecharon fuentes accesibles como *YouTube*, aplicaciones móviles, grabaciones de competencias y entrenamientos deportivos en actividades de equipo en deportes como voleibol, baloncesto, fútbol y tenis durante diversas situaciones. El estudio busca identificar a quien realiza la acción en una secuencia de video, determinar la duración de la acción y categorizar el tipo de acción para realizar análisis estadísticos automáticos. En técnicas de aprendizaje automático (ML) donde se utilizaron dos enfoques: extracción de características tradicional y aprendizaje profundo (DL). Se utilizaron varios métodos, como Máquinas de Soporte Vectorial (SVM), redes neuronales convolucionales (CNN), redes neuronales recurrentes (RNN)/LSTM, CNN 3D y modelos de atención/transformadores para analizar datos de video. La propuesta de sistematización de acciones humanas en deportes es crucial para categorizar y reconocer acciones, independientemente de las variaciones en la ejecución de estas acciones por parte de los jugadores.

Reconocimiento de movimiento de miembros superiores

A continuación, se proporcionará un contexto sólido y fundamentado en la ciencia de datos para el procesamiento de señales de miembros superiores. Se explorarán conceptos y técnicas claves aplicadas a esta área, con el fin de mejorar la comprensión y la aplicación de métodos efectivos en el análisis de datos de miembros superiores.

El artículo [47] proporciona información relevante sobre el uso de algoritmos de aprendizaje automático para el procesamiento de señales en el reconocimiento de movimientos del hombro. Se utilizaron señales de electromiografía de superficie (sEMG) en el reconocimiento del movimiento del hombro y se recolectan señales de múltiples músculos del hombro mientras los sujetos realizan movimientos de actividades cotidianas como beber, empujar hacia

II. ANTECEDENTES

adelante/tirar hacia atrás y abducción/aducción. Estas señales se procesan luego utilizando algoritmos de aprendizaje automático como máquina de vectores de soporte (SVM) con una función de kernel, regresión logística (LR) y una red neuronal artificial (ANN) que utiliza el algoritmo de retropropagación para ajustar los pesos entre los nodos en cada capa y minimizar el error entre las salidas deseadas y las reales para reconocer patrones en la actividad muscular que corresponden a movimientos específicos del hombro. El estudio comparó la eficiencia entre estas técnicas e investigó los efectos del intervalo de tiempo de la ventana deslizante en la precisión del reconocimiento. Como resultado, SVM tuvo la mejor precisión.

Un área de investigación que conlleva una relación con los miembros superiores [48], tiene como objetivo crear un sistema de control inteligente para la asistencia automática al operador de un exoesqueleto semi-pasivo. Se comparan dos enfoques para crear modelos humanos para el reconocimiento de acciones: uno basado en la reconstrucción de un modelo cinemático humano simplificado a partir de datos IMU y otro basado en la obtención de coordenadas 2D de los puntos clave de una persona a partir de video de dos vistas diferentes. Los conjuntos de modelos humanos obtenidos a partir de estos enfoques se clasifican utilizando el algoritmo XGBoost, y los clasificadores resultantes se comparan en términos de precisión. Se eligieron dos acciones para el reconocimiento: levantar una mano por encima de 90° con respecto al cuerpo humano y pintar una pared por una persona. Se creó una tabla para cada video que indicaba el número de fotograma y dos variables objetivo-binarias correspondientes a ese fotograma que debían predecirse en la tarea. Los valores de las variables objetivo se completaron manualmente al ver los datos obtenidos del video. Los resultados de entrenamiento obtenidos en la primera acción objetivo dieron como resultado una ROC AUC (*Receiver Operating Characteristic – Area Under the Curve*) del 98% y una precisión del 94% en los datos de video y una ROC AUC del 96% y una precisión del 93% en los datos de IMU. Para la segunda acción objetivo, se obtuvo una ROC AUC del 95% y una precisión del 84% en los datos de video y una ROC AUC del 98% y una precisión del 93% en los datos IMU. El estudio concluye que el enfoque basado en video supera al enfoque basado en IMU.

II. ANTECEDENTES

Con el objetivo de mapear de manera integral las sinergias musculares en una variedad de movimientos y sujetos, en el estudio redactado en el artículo [49], se considera la diversidad de movimientos, condiciones y entornos naturales que emulan la vida real. Se utilizó un algoritmo de factorización de matriz no negativa (NMF) para extraer las sinergias musculares. El proceso implicó la aplicación del algoritmo NMF, 50 veces para cada configuración y así evitar mínimos locales. Se agregaron sinergias adicionales solo si el total de la variación explicada aumentaba en al menos un 5% para cada sinergia adicional. Se consideraron tres ángulos relevantes: flexión del hombro, rotación vertical del hombro y flexión del codo. Los datos de 14 canales sEMG se filtraron en alta frecuencia a 50 Hz para eliminar artefactos de movimiento, se rectificaron y se filtraron en baja frecuencia con una frecuencia de corte de 10 Hz para extraer el sobre EMG. Los datos de cada tipo de movimiento se promediaron intra-sujeto para caracterizar un patrón medio, etiquetado como "EMG filtrado y promediado".

El artículo [50] se centra en cuantificar la fiabilidad de las métricas de calidad de movimiento en adultos sanos durante una variedad de actividades diarias no restringidas (ADLs). Diecinueve participantes (9 hombres con edades de 22 ± 4 años) sin problemas de salud, realizaron seis ADLs en dos sesiones separadas. Cada ADL se dividió en fases de alcance y manipulación de objetos. Se calculó la fiabilidad dentro y entre sesiones usando coeficientes de correlación intraclass (ICCs) y cambios detectables mínimos (MDC). El ICC superó el 0.75 para el 73% de las métricas entre sesiones, pero solo para el 5% de las métricas dentro de la sesión. El filtrado afectó la magnitud de las métricas de calidad del movimiento, especialmente para el *jerk* sin dimensiones logarítmico y el número de submovimientos. Se subraya la importancia de seleccionar adecuadamente las métricas para evaluaciones funcionales.

El artículo [51] busca mejorar el rendimiento en la decodificación de movimientos funcionales de una sola extremidad superior a partir de señales EEG de imaginación motora (MI) utilizando una red neuronal de wavelet (WNN). Se utilizó la base de datos de señales EEG pública de BNCI Horizon 2020, la cual tiene señales de 15 sujetos sanos de entre 22 y 40 años, quienes realizan seis movimientos funcionales de un solo miembro superior: pronación/supinación del antebrazo, apertura/cierre de la mano y flexión/extensión del codo.

II. ANTECEDENTES

Se utilizó la WNN para descomponer las señales EEG de MI en sub-bandas, de las cuales se extrajeron características estadísticas para codificar los movimientos funcionales. Posteriormente, se utilizó el análisis de componentes principales (PCA) para la selección del vector de características. La metodología incluyó el preprocesamiento de señales, segmentación de datos, extracción de características, estandarización, selección y clasificación basada en WNN. Como resultado, se logró la mejor precisión del $86.27 \pm 6.98\%$. Además, para verificar la eficiencia del clasificador WNN, se realizaron experimentos comparativos con la máquina de soporte vectorial (SVM), el análisis discriminante lineal (LDA), el vecino más cercano k (KNN) y la red neuronal de una sola capa oculta (ANN). Los resultados mostraron que la precisión del WNN mejoró aproximadamente un 15 ~ 40%, lo que demostró la efectividad del método WNN.

Se utilizó el aprendizaje automático en [52], donde se presenta una revisión de los avances en el campo de la ciencia del movimiento humano, el cual se centra en el análisis de movimiento en libertad. Se describen diferentes tipos de aprendizaje automático, incluyendo el aprendizaje supervisado, no supervisado y por refuerzo, los cuales se utilizan para revisar sus posibles aplicaciones en el análisis del rendimiento deportivo. El aprendizaje supervisado puede utilizarse para tareas de clasificación y regresión, así como distinguir entre patrones de movimientos normales y anormales o predecir medidas de rendimiento. El aprendizaje no supervisado puede utilizarse para agrupar y reducir la dimensionalidad, identificar grupos con técnicas análogas o entender similitudes entre los movimientos de los individuos. También, se discuten los desafíos de la selección y ajuste de modelos, incluyendo el gran número de modelos posibles y las diferentes tareas que se pueden realizar.

Para el uso de redes neuronales profundas, se realizan análisis cinemáticos sin marcadores en biomecánica [53]. Se requiere la combinación de una o más cámaras con un enfoque llamado estimación de pose para detectar puntos de referencia corporales y luego utilizar la geometría para estimar los ángulos de las articulaciones. El artículo discute los desafíos y limitaciones del uso de métodos de estimación de pose de aprendizaje profundo para el análisis cinemático en biomecánica y sugiere el uso de transferencia de aprendizaje y de etiquetado para mejorar la

II. ANTECEDENTES

precisión. Estas etiquetas indican dónde se encuentra una parte del cuerpo o un objeto en la imagen. Después de ver suficientes ejemplos, la red neuronal puede aprender a identificar esa parte del cuerpo en otras imágenes que no ha visto antes. La transferencia de aprendizaje permite que el conocimiento se transfiera entre tareas, lo que puede reducir la necesidad de grandes conjuntos de datos de entrenamiento.

III. ANÁLISIS DE DATOS DEL M²S

La creación y el diseño de bases de datos para investigaciones en áreas poco exploradas constituyen un pilar fundamental para el avance del conocimiento científico y tecnológico. En el ámbito del deporte, y más específicamente en el estudio del movimiento humano dentro del contexto del fútbol americano, es esencial desarrollar modelos de análisis y simulación que permitan una comprensión más profunda de las dinámicas del lanzamiento de balón. Por ello, este capítulo presenta la elaboración de la base de datos desarrollada en esta tesis y un análisis de la información contenida en esta.

La base de datos se diseñó para encontrar la forma más adecuada de visualizar el movimiento de una persona que realice un lanzamiento de balón de Fútbol Americano. Dicha base de datos se elaboró debido a que no se encontró información suficiente en el análisis del estado del arte realizado para la investigación en esta área. Ante esta necesidad, se ha concebido la base de datos de *Direct Pass Movement Dataset* (DPMov), una iniciativa pionera destinada a llenar este vacío.

El presente capítulo se organiza de la siguiente manera. La sección 3.1 presenta el protocolo para la elaboración de la base de datos, el cual considera el diseño del escenario y la toma de datos de cada uno de los participantes. La sección 3.2 describe el proceso contenido en los videos. La sección 3.3 presenta un análisis categórico basado en agrupamiento. Finalmente, la sección 3.4 presenta los resultados del análisis categórico.

3.1 Protocolo

Para diseñar la base de datos DPMov, se elaboró un protocolo que fue resultado de un riguroso proceso de diseño que tiene como objetivo garantizar la captura de información relevante del movimiento de pase directo de balón de Fútbol Americano. La definición de este protocolo abarcó el diseño del escenario, la selección de voluntarias y voluntarios y la especificación de los procedimientos de grabación de vídeo. Cada aspecto ha sido considerado con el fin de maximizar la relevancia de los datos recopilados.

3.1.1. Escenario

La adquisición de datos de DPMov implicó un meticuloso proceso de grabación de videos diseñado para capturar con precisión el movimiento de lanzamiento de balón de Fútbol Americano. Se cuidó que este proceso de grabación pueda ser replicado y se explicará a continuación.

El proceso se llevó a cabo utilizando un dispositivo móvil *Huawei Y9 Prime*, el cual se seleccionó por su capacidad para grabar a una resolución de 1080 pixeles. Esta resolución fue elegida porque es un estándar en dispositivos móviles y, por lo tanto, es accesible para cualquier persona que decida emplear este mismo protocolo. Además, esta resolución asegura que los detalles del movimiento de los sujetos fueran capturados de forma tal que se pueda realizar un análisis profundo.

Para garantizar la uniformidad y la estabilidad de las tomas, el celular fue montado en un trípode durante las sesiones de grabación. Este trípode no solo facilitó la captura de videos sin vibraciones ni movimientos indeseados, también permitió hacer ajustes precisos de altura. El trípode se ajustaba individualmente para cada persona, asegurando que el encuadre del video se mantuviera constante en términos de ángulo y perspectiva. Este ajuste era esencial para mantener el mismo campo visual en cada uno de los participantes, facilitando la comparación y el análisis de los movimientos de lanzamiento capturados.

En la Figura 3.1 se presenta la forma de colocar el trípode respecto a los participantes, quienes tienen diferencias en estatura. Se observa que se modifica la distancia del trípode al foco central y se ajusta el enfoque para tener modificar el campo de visión. Además, se ajusta la altura del trípode de la cámara para enfocarla en la parte superior del cuerpo y la distancia horizontal del trípode al sujeto para poder capturar el movimiento completo del brazo. Por ejemplo, para una persona de 1.85 metros, se ajusta el trípode a una altura de 1.57 metros y en distancia horizontal alrededor de 4.5 metros para tener la toma desde debajo de la cadera, hasta 50 centímetros por encima de la cabeza del voluntario. Por otro lado, en el caso de una persona que tenga una estatura baja, como un niño de 6 años, se ubica el trípode a una altura de 1.25

III. ANÁLISIS DE DATOS DEL M²S

metros y a una distancia de 3 metros, para ajustar el foco aproximadamente en las mismas circunstancias que la situación anterior.

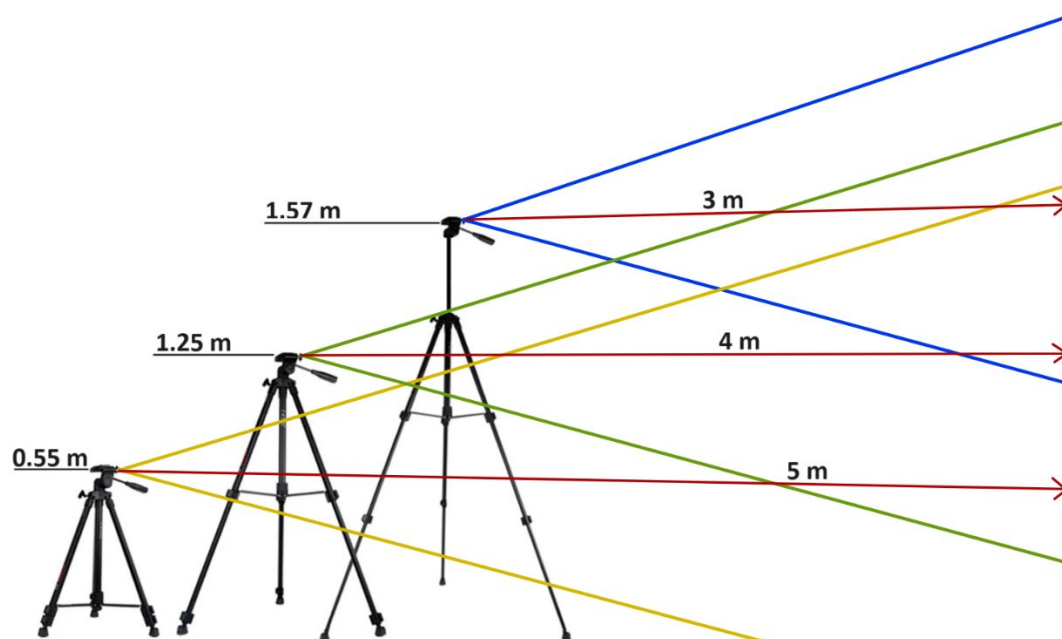


Figura 3.1. Diagrama de los cambios en altura y distancia del trípode.

Además del trípode y la cámara de 1080 pixeles, se tomó la decisión de utilizar una lona blanca como fondo durante la grabación de los videos. Esta lona tuvo varios propósitos:

- Proporcionaba un fondo neutro y uniforme que elimina las distracciones visuales, permitiendo que el foco de atención permaneciera en el sujeto y su movimiento.
- La uniformidad del fondo facilitaba la identificación y el seguimiento del balón y las extremidades del sujeto en las fases de análisis de video.
- La lona blanca mejoraba el contraste entre el sujeto y el fondo, asegurando que los detalles finos del movimiento se capturaran con claridad.

3.1.2 Toma de datos de voluntarias y voluntarios

El protocolo adoptó un enfoque inclusivo en la selección de sus participantes para asegurar una comprensión profunda y representativa de los movimientos asociados con el lanzamiento de un balón de Fútbol Americano. Este enfoque permitió el análisis comparativo de los patrones

III. ANÁLISIS DE DATOS DEL M²S

de movimiento entre distintos grupos, relacionados con sexo, edad, dimensiones corporales y experiencia en la técnica de lanzamiento.

Los voluntarios, dieron su consentimiento informado para participar en el estudio y llenaron un formato que pide datos de estatura, peso, sexo, si son jugadores con experiencia o no y el brazo dominante (izquierdo o derecho). Se considera como hipótesis en esta tesis que la variable que mide si el voluntario es jugador o no, mide el nivel de habilidad y familiaridad con el deporte. La Tabla 3.1 presenta el formato utilizado para recopilar los datos, el cual incluye edad (años), estatura (metros), peso (kilogramos), sexo (1 para masculino y 0 para femenino), jugador (1 para habitual y 0 para no habitual) y el brazo predominante para lanzar (0 para derecha y 1 para izquierda).

Esta información proporciona una referencia rápida para identificar patrones o correlaciones potenciales entre los atributos físicos de los sujetos y sus técnicas de lanzamiento. Dicha referencia servirá para definir las categorías y clases que se pueden obtener con DPMov.

3.2 Descripción de la base de datos

El proceso de captura de video para obtener datos sobre el lanzamiento de balón de fútbol americano se llevó a cabo implementando los elementos presentados en la sección 3.1. Este proceso se inició con la preparación del escenario, que incluyó la instalación de una lona blanca para proporcionar un fondo neutro y uniforme que minimizara las distracciones visuales y facilitara el seguimiento del movimiento del sujeto.

Una vez que se estableció el escenario, se posicionó el trípode que sostenía el celular de acuerdo con el protocolo establecido (a una distancia entre 3 y 5 metros del sujeto). Se instruyó a las voluntarias y los voluntarios, tanto a aquellos que son frecuentes lanzadores de balón como aquellos que lo hacen de forma infrecuente para que realizaran un lanzamiento hacia una distancia aproximada de 5 metros y de forma directa. Esta distancia se seleccionó por su idoneidad para capturar el movimiento de lanzamiento en su totalidad, permitiendo una evaluación exhaustiva de la técnica y la biomecánica involucradas en el acto de lanzar el balón. Mediante este protocolo, se obtuvieron 106 videos de 38 sujetos.

III. ANÁLISIS DE DATOS DEL M²S

Tabla 3.1. Tabla de documentación que proporciona una muestra representativa de datos de los individuos tomados en video.

Video	Edad	Estatura	Peso	Genero	Jugador	Percepción si es profesional	Der0/izq1
1	29	1.59	56	1	1	1	0
2	31	1.61	62	0	1	1	0
3	12	1.56	69	1	0	0	0
4	19	1.75	78	1	1	0	1
5	28	1.88	84	1	1	0	0
6	35	1.64	63	0	0	0	0
7	26	1.68	76	1	1	1	1
8	8	1.41	26	0	0	0	0
9	42	1.71	86	1	0	0	0
10	9	1.3	24	1	0	0	0
11	6	1.22	21	1	0	0	0
12	34	1.6	52	0	1	0	0
13	39	1.63	64	1	1	0	0
14	56	1.76	88	1	0	0	0
15	53	1.58	61	0	0	0	0
16	29	1.68	76	0	1	1	1

En la siguiente sección, se describen las cuatro posiciones clave del movimiento del brazo según la técnica estandarizada por Dub Maddox [54], destacando su importancia en el lanzamiento y su impacto en la técnica. Además, se analizan las diferencias en las técnicas de lanzamiento entre jugadores frecuentes e infrecuentes, subrayando cómo la experiencia y la práctica influyen en la eficiencia y eficacia del lanzamiento.

3.2.1 Lanzamiento directo de balón

El entrenador y miembro de *Quarterback Academy*, Dub Maddox, reportó en [54] que el lanzamiento de balón es la habilidad motora más compleja en todos los deportes. Por lo que estandarizó una técnica que consta de 4 posiciones clave del movimiento del brazo en la cadena cinética. Este estándar se basó en trabajo en expertos del deporte, kinesiólogos, médicos y consiste en los siguientes pasos:

1. Se comienza por la posición llamada “Triangulo Pre-Pase”, que se refiere a la posición del brazo antes de lanzar el balón, como se muestra en la Figura 3.2. Consiste en mantener los codos a la misma altura y la muñeca en una posición cargada detrás del hombro del brazo no lanzador. Esta configuración crea una forma triangular que proporciona una base sólida y potente para lanzar el balón con fuerza. Además, esta posición facilita la carga

III. ANÁLISIS DE DATOS DEL M²S

del brazo para lanzar el balón de forma instintiva. El Triángulo Pre-Pase ayuda a evitar la rotación interna del brazo durante el lanzamiento, asegurando una alineación adecuada y reduciendo la pérdida de movimiento.



Figura 3.2. Muestra del Triángulo Pre-Pase hecha por jugadores profesionales de fútbol americano.

2. La Transición en forma de "L" es la siguiente etapa en el proceso de lanzamiento. Para llegar a esta posición, se activan los 4 músculos del manguito rotador que rodean la escápula, como se observa en la Figura 3.3. Dos de estos músculos, el infraespinoso y el redondo menor rotan externamente el brazo para llevarlo a una posición en forma de "L". Cuando el brazo está en esta posición, se estiran el músculo supraespinoso y el subescapular, lo que permite a los demás músculos acelerar el codo hacia adelante en preparación para el lanzamiento.

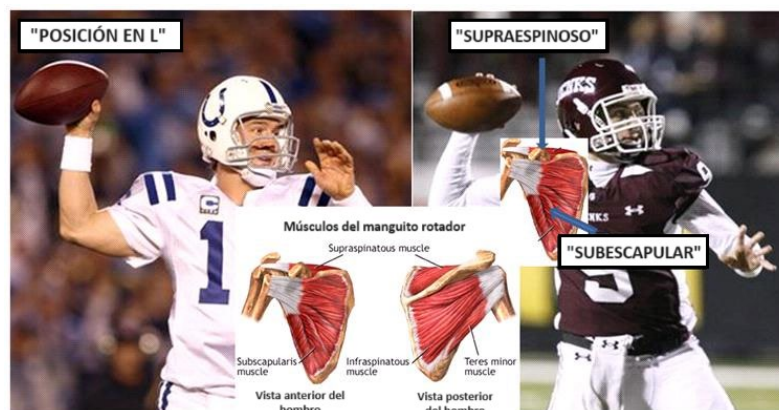


Figura 3.3. Posición en forma de "L" hecha por jugadores profesionales y los músculos asociados.

III. ANÁLISIS DE DATOS DEL M²S

3. El siguiente paso es elevar a la posición de "Cero", el cual es cuando el codo está en la posición principal para ayudar a estabilizar la muñeca durante el lanzamiento, como se muestra en la Figura 3.4. Después de cargar los músculos supraespinoso y subescapular en la posición en forma de "L", el codo puede elevarse y moverse adelante del hombro con la ayuda del deltoides para llegar a la posición de "Cero". Se llama "Cero" porque los músculos del manguito rotador están en una posición neutral y sin tensión. La posición de "Cero" coloca el codo 6 pulgadas adelante del hombro, en un ángulo de 45 grados hacia arriba y hacia afuera (135° respecto a la horizontal), preparando el tríceps para impulsar la pelota hacia el objetivo.



Figura 3.4. Posición "Cero" de los jugadores profesionales.

4. El último paso es la "Extensión", es el proceso en el que el tríceps libera energía a lo largo del brazo, transmitiéndola hasta la muñeca y los dedos, para finalmente transferirla al balón durante el lanzamiento. Si la muñeca se mueve antes que el tríceps haga su trabajo, la secuencia de energía se desorganiza y la pelota no irá donde se espera. Si un mariscal de campo no estira completamente el brazo al lanzar, no estará utilizando toda la fuerza del tríceps. En la Figura 3.5 se puede observar este movimiento.



Figura 3.5. Extensión del brazo en los jugadores profesionales de fútbol americano.

3.2.2 Lanzamiento del balón en voluntarios

Para esta investigación en particular, cada voluntario tuvo la libertad de proporcionar su mejor técnica de lanzamiento, lo que permitió capturar una variedad de estilos y enfoques en los videos recopilados. Esta diversidad en las técnicas de lanzamiento derivada de la variabilidad en la experiencia y la habilidad de los participantes, enriqueció la base de datos DPMov, ya que se pudo observar una amplia gama de información que reflejaban la realidad del movimiento de lanzamiento de pase directo en el Fútbol Americano.

A partir de la toma de estos videos, se visualizan diferencias a simple vista de varios aspectos relacionados con el movimiento y la técnica de lanzamiento entre jugadores frecuentes e infrecuentes. A continuación, se presentan algunas posibles diferencias que podrían existir entre estos dos grupos.

Jugadores frecuentes.

1. **Experiencia y habilidad técnica.** Exhiben una técnica de lanzamiento que ha sido perfeccionada y optimizada a través de la práctica continua y experiencia en el deporte.
2. **Desarrollo de fuerza muscular.** El entrenamiento específico ha permitido desarrollar una mayor fuerza en los músculos relevantes para el lanzamiento, como los del brazo, hombros y torso.
3. **Confianza y actitud mental.** Poseen una mayor confianza en sus habilidades de lanzamiento y una actitud mental más positiva hacia el desafío de lanzar.

4. Riesgo de lesiones. Están más expuestos al riesgo de lesiones por sobreuso debido a la frecuencia con la que practican el lanzamiento.

Jugadores infrecuentes

1. Experiencia y habilidad técnica. Es probable que exhiban una técnica de lanzamiento menos fluida o controlada debido a su falta de práctica regular y experiencia limitada en el deporte.
2. Desarrollo de fuerza muscular. Pueden tener una fuerza menos desarrollada en los músculos relevantes para el lanzamiento debido a la falta de entrenamiento específico.
3. Confianza y actitud mental. Es posible que tengan una menor confianza en sus habilidades de lanzamiento y una actitud mental menos positiva debido a su falta de experiencia, lo que podría afectar su rendimiento durante las situaciones de lanzamiento.
4. Fatiga muscular. Son más susceptibles a la fatiga muscular debido a la falta de acondicionamiento específico para el lanzamiento y la falta de práctica regular.

Estas son solo algunas de las posibles diferencias que podrían existir entre las personas que son frecuentes e infrecuentes lanzadoras y lanzadores de balón de Fútbol Americano. Es importante tener en cuenta que la variabilidad entre individuos puede ser considerable y que estas diferencias pueden ser influenciadas por una variedad de factores, incluyendo la edad, el género, la condición física general y la historia deportiva de cada individuo.

3.3 Análisis categórico

El análisis categórico es un conjunto de técnicas estadísticas que se utilizan para analizar variables cualitativas o categóricas. Las variables categóricas son aquellas que describen atributos o cualidades y no tienen un orden inherente. Tiene una amplia gama de aplicaciones en diversos campos debido a su capacidad para agrupar datos en categorías significativas.

Para el caso de esta tesis, se utiliza el análisis categórico para conocer si la variable de jugador/no-jugador se pueden utilizar como clases para clasificación. La metodología del análisis categórico de esta tesis está inspirada en el método propuesto en [55] y consiste en

III. ANÁLISIS DE DATOS DEL M²S

tomar la información antropométrica y estadística de los participantes para generar variables que se puedan aplicar a modelos de *clustering*. Posteriormente, la salida de los modelos de *clustering* será comparada con las categorías de la variable jugador/no-jugador.

La hipótesis de este análisis categórico es que los videos de la base de datos DPMov se pueden separar en dos categorías basadas en videos de jugador infrecuente y jugador frecuente. La categoría jugador infrecuente se refiere a aquellos que participan en el deporte de manera esporádica o con poca frecuencia. Estos jugadores pueden tener menos experiencia y habilidad técnica debido a su falta de práctica regular y su participación limitada en el deporte. La categoría jugador frecuente se refiere a aquellos jugadores que participan regularmente en el deporte, practicando y compitiendo con regularidad. Estos jugadores dedican mucho tiempo y esfuerzo al entrenamiento y al juego, lo que les brinda una mayor experiencia y habilidad técnica.

Para el análisis categórico, es necesario revisar que variables demográficas y estadísticas son relevantes en la clasificación de jugador infrecuente y jugador frecuente. Las variables consideradas en esta tesis se obtienen a partir de los estatutos físicos proporcionados por los mismos participantes, como la estatura, peso, sexo, edad, brazo lanzador e información de los videos captados del lanzamiento. Estas variables fueron seleccionadas ya que proporcionan información importante sobre las características físicas y técnicas de los participantes, que pueden influir en su rendimiento en el deporte. A continuación, se presentan las variables seleccionadas:

Edad (V_1). La edad puede afectar la capacidad física y técnica de un individuo. Las personas de diferentes edades pueden tener diferentes niveles de fuerza, flexibilidad y coordinación, lo que podría influir en su técnica de lanzamiento. Además, el desarrollo físico y técnico puede variar a lo largo de la vida de una persona.

Estatura (V_2). La estatura puede afectar la mecánica del lanzamiento, incluyendo la longitud del brazo y la altura desde la que se lanza el balón. Las personas más altas pueden tener un alcance mayor, pero deben ajustar su técnica de lanzamiento en consecuencia.

Peso (V_3). El peso corporal puede influir en la fuerza y la estabilidad del individuo durante el lanzamiento. Las personas con un peso mayor pueden tener más fuerza detrás de su lanzamiento, pero también pueden requerir un mayor control de su movimiento para mantener la estabilidad durante el lanzamiento.

Sexo (V_4). Hay diferencias biológicas entre hombres y mujeres que pueden influir en la técnica de lanzamiento. Por ejemplo, los hombres tienden a tener más fuerza muscular y pueden lanzar el balón con más potencia, mientras que las mujeres pueden tener una mayor flexibilidad en las articulaciones que podría afectar su técnica de lanzamiento.

Izquierdo o derecho (V_5). La adaptación técnica y estratégica varía significativamente entre lanzar con el brazo izquierdo o derecho, afectando la mecánica del lanzador.

Varianza del cambio de cuadro por *Gunnar Farnebäck* ($V_6, \dots, V_9$). Se refiere a la varianza de los niveles de grises dados por:

$$\sigma_j^2 = \frac{1}{n-1} \sum_{i=1}^n (D_j - \mu_{Dj})^2 \quad (3.1)$$

Donde $D_j = g[|I^t - I^{t-1}|]$ y μ_{Dj} es la media de los niveles de grises de D_j , $g[u]$ se refiere a la diferencia espacial de los momentos centrales de primero ($j=1$), segundo ($j=2$), tercer ($j=3$) y cuarto orden ($j=4$) de *Gunnar Farnebäck*. La varianza σ_j^2 contiene información del movimiento generado por el sujeto.

Esta variable se analiza considerando un video por sujeto.

Experiencia como jugador habitual (w_1). Variable categórica que se asemeja más a jugador frecuente o infrecuente. Los jugadores habituales pueden haber desarrollado habilidades y técnicas específicas a lo largo del tiempo, mientras que los no habituales pueden tener menos experiencia y estar menos familiarizados con la mecánica adecuada del lanzamiento.

Estas variables contienen la información general para el análisis categórico. Una vez que se identificaron las variables del experimento del análisis categórico, se debe realizar un análisis por coeficiente de correlación *Pearson* y Análisis de Componentes Principales (PCA) para

identificar la relación estadística en estas variables y definir las características del análisis categórico. A continuación, se describirán la correlación de *Pearson* y PCA.

3.3.1 Coeficiente de correlación de *Pearson*

El coeficiente de correlación de *Pearson* es una medida estadística que cuantifica la fuerza y la dirección de una relación lineal entre dos variables continuas. Este coeficiente se representa comúnmente por la letra "*r*" y tiene un valor que varía entre -1 y 1. La fórmula para calcular el coeficiente de correlación de *Pearson* es:

$$r = \frac{\sum(V_x - \mu_{V_x})(V_y - \mu_{V_y})}{\sqrt{\sum(V_x - \mu_{V_x})^2 \sum(V_y - \mu_{V_y})^2}} \quad (3.2)$$

Donde V_x y V_y son las variables utilizadas y μ_{V_x} y μ_{V_y} son sus medias.

La interpretación de valores se expresa de la siguiente manera:

$r=1$: indica una correlación lineal positiva perfecta, donde un aumento en una variable está perfectamente asociado con un aumento en la otra variable.

$r=-1$: indica una correlación lineal negativa perfecta, donde un aumento en una variable está perfectamente asociado con una disminución en la otra variable.

$r=0$: indica que no hay correlación lineal entre las dos variables.

Las características del Coeficiente de Correlación de *Pearson* son:

Dirección: el signo del coeficiente indica la dirección de la relación (positivo para una relación directa, negativo para una relación inversa).

Magnitud: el valor absoluto del coeficiente indica la fuerza de la relación (valores más cercanos a 1 o -1 indican una relación más fuerte).

Linealidad: *Pearson* mide solo relaciones lineales. Si la relación entre las variables es no lineal, el coeficiente de *Pearson* no será una representación adecuada de la relación.

La Tabla 3.2 presenta la matriz de correlación donde la entrada son las variables de $V = \{V_1, V_2, V_3, V_4, V_5, V_6\}$. Se puede observar que V_2 y V_3 tienen una correlación alta, las cuales son

III. ANÁLISIS DE DATOS DEL M²S

estatura y peso, respectivamente. La variable V_4 (sexo) presenta una buena correlación con V_2 y V_3 . Las variables V_7 (segundo momento central) y V_9 (tercer momento central) presentan una correlación muy alta. Basándose en los resultados de la correlación y el concepto de las variables, se puede asumir lo siguiente:

- El peso y estatura de los sujetos de la base de datos de DPMov se puede resumir en una variable con métodos como PCA.
- El segundo y cuarto momento central de *Gunnar-Farneback* se puede resumir en una variable con métodos como PCA.
- El sexo se relaciona con estatura y peso, ya que los hombres suelen tener una mayor estatura y peso, mientras que las mujeres tienden a presentar menor estatura y peso. Dado que esta información no se relaciona con la técnica y el movimiento muscular, no se consideró para resumir V_2 , V_3 y V_4 en una variable.⁴

Tabla 3.2. Matriz de correlación.

	V_1	V_2	V_3	V_4	V_5	V_6	V_7	V_8	V_9
V_1	1	0.1747	0.2032	0.2419	-0.1642	-0.023	-0.3081	-0.25093	-0.31253
V_2	0.17471	1	0.8779	0.7334	-0.0416	0.0679	-0.0274	-0.1468	-0.11694
V_3	0.20325	0.8779	1	0.7429	-0.0717	0.1753	0.0966	-0.17186	0.00188
V_4	0.2419	0.7334	0.7429	1	-0.1197	0.2953	0.285	0.078217	0.23114
V_5	-0.1642	-0.042	-0.072	-0.12	1	-0.1528	-0.1735	0.21339	-0.08607
V_6	-0.023	0.0679	0.1753	0.2953	-0.1528	1	0.6377	0.475229	0.66355
V_7	-0.3081	-0.027	0.0966	0.285	-0.1735	0.6377	1	0.439386	0.94554
V_8	-0.2509	-0.147	-0.172	0.0782	0.21339	0.4752	0.4394	1	0.58669
V_9	-0.3125	-0.117	0.0019	0.2311	-0.0861	0.6635	0.9455	0.586693	1

3.3.2 Análisis de Componentes Principales

El Análisis de Componentes Principales (PCA, por sus siglas en inglés) es un modelo estadístico multivariable utilizado para reducir la dimensionalidad del conjunto de datos, pero preservando la varianza de estos. En esencia, se buscan las direcciones (o componentes principales) en las que los datos varían más. El proceso de PCA implica los siguientes pasos:

1. Estandarización de los datos: todas las variables se normalizan para garantizar que tengan la misma escala de la siguiente manera:

$$Z_x = \frac{V_x - \mu_{V_x}}{\sigma_{V_x}} \quad (3.3)$$

Donde Z_x es el valor estandarizado, V_x es el valor original de la variable, μ_{V_x} es la media de la variable y σ_{V_x} es la desviación estándar de la variable.

Esta transformación ajusta los datos de manera que cada variable tenga una media de 0 y una desviación estándar de 1. Esta estandarización permite comparar variables en diferentes escalas y magnitudes de manera equitativa, asegurando que ninguna variable domine el análisis debido a su rango o unidades.

2. Cálculo de la matriz de covarianza: a partir de los datos estandarizados, se calcula la matriz de covarianza que describe las relaciones lineales entre las variables. Un valor positivo de covarianza indica que cuando una variable aumenta, la otra también tiende a aumentar, mientras que un valor negativo indica una relación inversa. La covarianza de los datos se define como:

$$C = Cov(Z_x, Z_y) = \frac{1}{n} \sum_{i=1}^n (Z_x - \mu_{Z_x})(Z_y - \mu_{Z_y}) \quad (3.4)$$

Donde $Cov(Z_x, Z_y)$ es la covarianza entre las variables normalizadas Z_x y Z_y , Z_x e Z_y , μ_{Z_x} y μ_{Z_y} son las medias de las variables Z_x y Z_y , respectivamente y n es el número de observaciones.

3. Descomposición de la matriz de covarianza: se utiliza la descomposición de valores propios (eigenvalores) y vectores propios (eigenvectores) para encontrar las direcciones principales de variabilidad en los datos. Los valores propios representan la cantidad de varianza explicada por cada componente principal, mientras que los vectores propios definen la dirección de cada componente principal en el espacio de las variables originales. La descomposición de la matriz de covarianza C se realiza de la siguiente manera:

$$C v_i = \lambda_i v_i \quad (3.5)$$

Donde C es la matriz de covarianza, λ_i son los valores propios (eigenvalores) y v_i son los vectores propios (eigenvectores).

En términos matriciales, la descomposición se puede expresar como:

$$C = U\Lambda U^T \quad (3.6)$$

Donde U es la matriz de vectores propios (cada columna de U es un vector propio v_i), Λ es la matriz diagonal de valores propios (los elementos diagonales de Λ son los valores propios λ_i) y U^T es la transpuesta de la matriz de vectores propios.

Los eigenvectores v_i se definen como los vectores no nulos que satisfacen la ecuación de eigenvalor:

$$(C - \lambda_i Id)v_i = 0 \quad (3.7)$$

Donde Id es la matriz identidad de la misma dimensión que C .

4. Selección de componentes principales: se ordenan los valores propios de mayor a menor de la siguiente manera:

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_p$$

Luego, se seleccionan los primeros k valores propios y sus correspondientes vectores propios (v_i) que representan la mayor parte de la variabilidad en los datos. El número k se elige de manera que la suma de los valores propios seleccionados represente un porcentaje significativo de la varianza total. Generalmente, se seleccionan los componentes principales que cubren al menos el 70-90% de la varianza total. Esto se puede representar como:

$$\sum_{i=1}^k \lambda_i \approx \text{un porcentaje significativo de la varianza total.}$$

Posteriormente, se construye la matriz de componentes principales (W) utilizando los primeros k vectores propios ordenados.

$$W = [v_1 \ v_2 \ v_3 \ \dots \ v_k]$$

5. Transformación de los datos: finalmente, se transforman los datos originales en un nuevo espacio de características definido por los componentes principales seleccionados. Esta transformación permite representar los datos en un espacio de menor dimensionalidad, donde las variables están altamente correlacionadas y la mayoría de la variabilidad se captura en un número reducido de dimensiones. Esta transformación se define por:

$$PCA = Z W \quad (3.8)$$

Donde Z_q son los datos transformados en el espacio de los PCA y W es la matriz de PCA formada por los vectores propios seleccionados.

Los componentes principales seleccionados (*PCA*) son combinaciones lineales de las variables originales V que capturan la mayor parte de la varianza en los datos. Estas combinaciones permiten una reducción de dimensionalidad efectiva sin perder información crucial sobre la estructura del conjunto de datos original.

Este análisis se aplicó con las variables V_2 y V_3 , obteniéndose un espacio nuevo de dos dimensiones $PCA = \{PCA_1, PCA_2\}$. La varianza de los datos de PCA_1 fue de 482.72, mientras que la varianza de PCA_2 fue de 0.004. Por lo tanto, se considera que PCA_1 describe la información de estatura y peso en una característica. También, se aplicó con las variables V_7 y V_9 , obteniéndose un espacio nuevo de dos dimensiones $PCA = \{PCA_3, PCA_4\}$. La varianza de los datos de PCA_3 fue de 0.28, mientras que la varianza de PCA_4 fue de 0.001. Por lo tanto, se considera que PCA_3 describe la información del primer y cuarto momento central en un vector de características.

Después de aplicar el PCA, se obtiene un vector para análisis categórico definido como: $\Gamma = \{V_1, PCA_1, V_4, V_5, V_7, V_8, PCA_3\}$. Este vector es la entrada de los algoritmos de análisis categórico basado en *clustering*. Para el caso de esta tesis, dichos algoritmos son K-Means y *Fuzzy C Means*, los cuales se describirán a continuación.

3.3.3 *Clustering* no Supervisado con *K-Means*

El *clustering* no supervisado es una técnica utilizada en el aprendizaje automático que agrupa datos similares en categorías o "*clusters*" sin que haya etiquetas predefinidas. Uno de los algoritmos más populares de *clustering* es el *K-Means*, que se utiliza para dividir un conjunto de datos en un número específico de grupos basados en la similitud entre los puntos de datos.

K-Means agrupa un conjunto de datos en un número K predefinido de grupos, donde cada dato pertenece al grupo con el centroide de cada grupo más cercano. El proceso de *K-Means* puede describirse en los siguientes pasos:

1. Inicialización de los centroides. Se inicializan K centroides como un conjunto de vectores con valores aleatorios.

2. Asignación de datos a *clusters*. Para cada dato se calcula la distancia euclidiana entre este y cada uno de los centroides. La distancia euclidiana d entre un punto de datos x_i y un centroide C_j se define como:

$$d(x_i, C_j) = \sqrt{\sum_{k=1}^n (x_{ik} - C_{jk})^2} \quad (3.9)$$

Donde x_i es el i -ésimo punto de datos, C_j es el j -ésimo centroide, x_{ik} es la k -ésima coordenada del i -ésimo punto de datos, C_{jk} es la k -ésima coordenada del j -ésimo centroide y n es la dimensión de los datos.

Para cada punto de datos x_i , calcular la distancia euclidiana a cada uno de los centroides C_j . Posteriormente, el grupo de cada i -ésimo dato x_i se asigna con la distancia euclidiana mínima dada por:

$$j = \arg \min_j d(x_i, C_j) \quad (3.10)$$

Este proceso se repite para todos los puntos de datos en el conjunto, asegurando que cada punto sea asignado al *cluster* más cercano basado en la distancia euclidiana a los centroides.

3. Actualización de los centroides. Una vez que todos los puntos han sido asignados a *clusters*, se vuelve a calcular el nuevo centroide C_j de la siguiente manera:

$$C_j = \frac{1}{N_j} \sum_{i=1}^{N_j} x_i \quad (3.11)$$

III. ANÁLISIS DE DATOS DEL M²S

Donde C_j es el nuevo centroide del *cluster* j , N_j es el número de puntos asignados al *cluster* j , x_i son los puntos de datos asignados al *cluster* j y la sumatoria se realiza sobre todos los puntos x_i pertenecientes al *cluster* j .

Para cada *cluster* j , se calcula la media dada por (3.11), la cual se establece esta media como el nuevo centroide C_j del *cluster*.

Este proceso de cálculo de centroides asegura que cada uno se reposicione en el centro geométrico de los puntos asignados a su *cluster*, reflejando la distribución actualizada de los datos.

4. Repetición de pasos 2 y 3. Los pasos 2 y 3 se repiten iterativamente hasta que se cumple un criterio de convergencia.

5. Convergencia y resultado final. Cuando se alcanza la convergencia se obtiene el resultado final. La convergencia se alcanza cuando los centroides no cambian significativamente entre iteraciones consecutivas. Esto se define como:

$$\max_j \|C_j^{(t)} - C_j^{(t-1)}\| < \epsilon \quad (3.12)$$

Donde $C_j^{(t)}$ es el centroide del *cluster* j en la iteración t , $C_j^{(t-1)}$ es el centroide del *cluster* j en la iteración $t-1$, $\|\cdot\|$ denota la norma euclidiana (distancia) y ϵ es un umbral pequeño predefinido que determina la tolerancia para el cambio en los centroides.

Si la variación máxima en la posición de los centroides es menor que el umbral ϵ , se considera que el algoritmo ha convergido.

3.3.4 Fuzzy C-Means (FCM)

Fuzzy C-Means (FCM) es un algoritmo de *clustering* que se basa en que los datos pueden pertenecer a más de un *cluster*, pero con diferentes grados de pertenencia. Es una técnica utilizada en análisis y minería de datos para identificar patrones y estructuras escondidas dentro de datos con altas dimensiones. A diferencia de métodos como *K-means*, donde cada dato pertenece estrictamente a un solo *cluster*, FCM proporciona una mayor flexibilidad y capacidad

para manejar la incertidumbre y la ambigüedad en la asignación de puntos a *clusters*. La función objetivo que FCM minimiza es:

$$J_m = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^m \|x_i - c_j\|^2 \quad (3.13)$$

Donde N es el número de puntos de datos, C es el número de *clusters*, m es el coeficiente de fuzzificación ($m > 1$), que determina el nivel de pertenencia, x_i es el i -ésimo punto de datos, u_{ij} es el grado de pertenencia del i -ésimo punto de datos al j -ésimo *cluster*, c_j es el centro del j -ésimo *cluster* y $\|x_i - c_j\|$ es la distancia entre el punto de datos x_i y el centro del *cluster* c_j . Los pasos para el algoritmo de FCM son:

1. Inicialización. Elegir el número de *clusters* C e inicializar la matriz de pertenencia U aleatoriamente, donde $u_{ij} \in [0,1]$ y $\sum_{j=1}^C u_{ij} = 1$ para todos los i -ésimos datos.

2. Actualización de los centros de los *clusters*. Calcular los nuevos centros de los *clusters* c_j usando la formula:

$$c_j = \frac{\sum_{i=1}^N u_{ij}^m x_i}{\sum_{i=1}^N u_{ij}^m} \quad (3.14)$$

3. Actualización de la matriz de pertenencia. Actualizar los grados de pertenencia u_{ij} usando:

$$u_{ij} = \frac{1}{\sum_{k=1}^C \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}} \quad (3.15)$$

4. Criterio de convergencia. Se repiten los pasos 2 y 3 hasta que los cambios en los centros de los *clusters* o en la matriz de pertenencia sean menores que un umbral definido.

3.4 Resultados del análisis categórico

El análisis estadístico y el *clustering* son herramientas fundamentales en el campo de la ciencia de datos y el aprendizaje automático. Estas técnicas permiten identificar patrones y relaciones dentro de conjuntos de datos complejos. En este contexto, los algoritmos *K-means* y

Fuzzy C-means (FCM) juegan un papel crucial en la segmentación de datos y la clasificación de patrones de movimiento, como los observados en estudios de movimiento humano.

3.4.1 Datos estadísticos de jugador o no jugador

Los algoritmos *K-means* y FCM se utilizaron en este análisis categórico para conocer cómo se agrupan los participantes y comparar ese agrupamiento con la clase w_1 (jugador vs no-jugador).

La clase w_1 tiene valores de uno para jugador y cero para no jugador. La forma en que se definió si un voluntario tiene el valor de cero o uno fue mediante el protocolo, donde a cada sujeto se le preguntó si jugaba de forma habitual al fútbol americano o no. Para visualizar y contrastar las diferentes consideraciones de los participantes, se creó la Figura 3.6 utilizando herramientas como Python (*matplotlib* y *seaborn*).

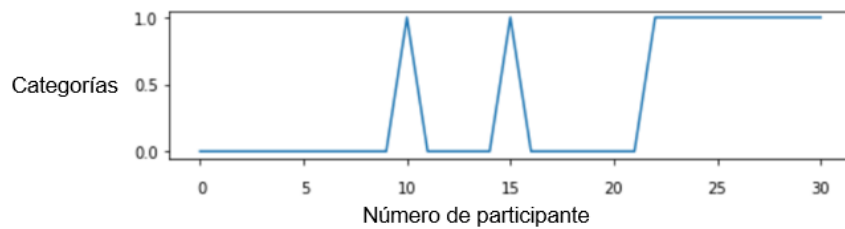


Figura 3.6. Visualización de la estadística que se perciben como jugadores habituales (1) y no habituales (0), en 30 participantes.

Esta gráfica ayudó a identificar patrones y tendencias en los datos. De esta forma, se puede comparar w_1 con los resultados que se obtienen en el *K-Means* y FCM.

3.4.2 Resultados de K-Means

La entrada del algoritmo *K-Means* para el análisis categórico es un vector que se define como $\Gamma = \{V_1, PCA_1, V_4, V_5, V_7, V_8, PCA_3\}$. Se normalizaron las características de Γ para asegurar que todas las variables tuvieran el mismo peso en el análisis. El *K-Means* tiene una cantidad de *clusters* de $K = 2$ y la salida se puede ver en la Figura 3.7.

III. ANÁLISIS DE DATOS DEL M²S

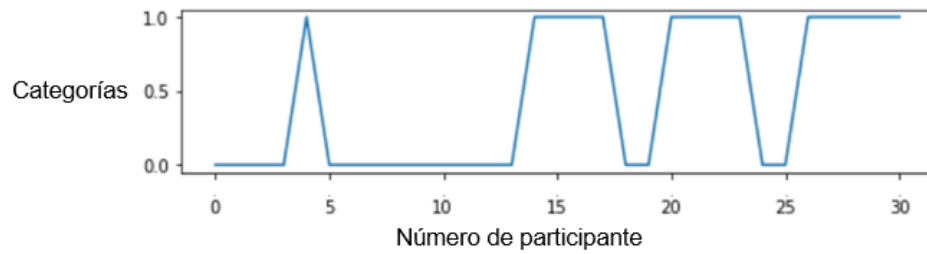


Figura 3.7. Visualización del análisis categórico con *K-means*.

Para este análisis de M²S, se eligió $K = 2$ para diferenciar entre puntos de datos de participantes que mostraron experiencia en el lanzamiento de un balón de futbol americano (unos) y participantes sin experiencia (ceros).

1. **Falso positivo de *K-Means*.** Se refiere a un video de un voluntario (el número 4) que *K-Means* lo marca como jugador, pero en w_I este se declaró no jugador. De acuerdo con el análisis realizado, se debió a problemas de iluminación que causaba errores en la medición del movimiento.
2. **Falsos negativos de *K-Means*.** Se refiere a videos de voluntarios que *K-Means* los marca como no jugador, pero en w_I se declararon jugadores. Estos fueron casos interesantes ya que fueron sujetos que juegan de manera frecuente, pero en la entrevista del protocolo manifestaron que se han lesionado de forma constante. Por lo tanto, es posible que *K-Means* los esté agrupando como técnica de movimiento incorrecta.
3. **Verdaderos positivos y negativos.** Los sujetos restantes son consistentes con los eventos esperados en los datos. De los 35 voluntarios, 26 fueron verdaderos positivos o negativos (74.2% de los sujetos).

3.4.3 Resultados de FCM

La entrada del algoritmo FCM para el análisis categórico es el vector $\Gamma = \{V_1, PCA_1, V_4, V_5, V_7, V_8, PCA_3\}$. Para implementar el FCM, se normalizaron las características de Γ para asegurar que todas las variables tuvieran el mismo peso en el análisis. Luego, se

III. ANÁLISIS DE DATOS DEL M²S

colocó una cantidad de *clusters* igual a 2 (al igual que en el *K-Means*), se corrió el FCM y la salida se puede ver en la Figura 3.8.

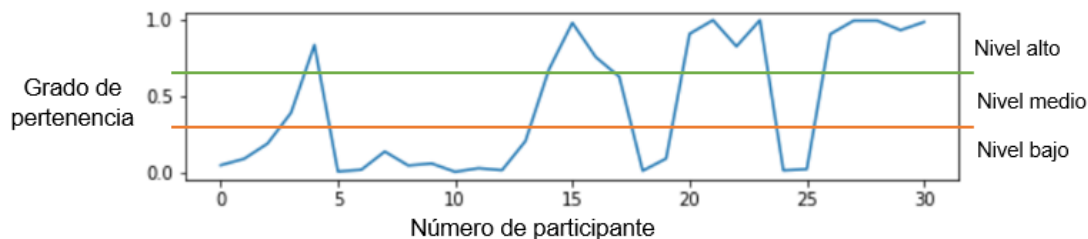


Figura 3.8. Visualización del análisis categórico con FCM. Se indica el nivel de pertenencia hacia jugador frecuente (1) e infrecuente (0).

Los grados de pertenencia obtenidos para cada sujeto en FCM proporcionaron información sobre la correlación entre estos y la técnica de lanzamiento. Por lo tanto, se analizaron estos grados de pertenencia para identificar patrones y relaciones entre las dos clases.

Dentro del análisis se observó que en el análisis de M²S en Fútbol Americano, los patrones de movimiento pueden no ser estrictamente diferenciables entre jugadores y no jugadores. Por lo tanto, este análisis se basó en visualizar la distribución de los grados de pertenencia para cada voluntario. Esto permitió identificar aquellos que tenían características de movimiento similares a las clases de w_1 y detectar jugadores con técnica intermedia, proporcionando una visión más matizada del rendimiento. A continuación, se presentan cuatro casos obtenidos de los resultados del FCM.

- Al igual que en el análisis de *K-Means*, el voluntario número 4 tiene un alto nivel de pertenencia al *cluster* asociado a la categoría jugador. Sin embargo, ya se había identificado previamente que se debía a un problema de iluminación, por lo cual se imposibilita reconocer el grado de pertenencia a cada categoría.
- Los participantes cuyos resultados fueron entre 0 y 0.35, se les podría catalogar como personas que tienen poca o nula experiencia en una correcta técnica de lanzamiento de Fútbol Americano.

III. ANÁLISIS DE DATOS DEL M²S

- En el caso de los individuos que se encuentran entre 0.36 y 0.65, se considera que están en un nivel medio entre ambas categorías. Es decir, se percibe que tienen cierto grado de experiencia, pero no tienen una buena técnica en el lanzamiento del balón.
- Para los participantes que tienen entre 0.66 a 1, se les podría considerar como personas que tienen una buena habilidad o técnica adecuada conforme lo indican los resultados estadísticos.

No obstante, estos casos son solo una estimación de lo que podrían significar los indicadores de cada uno de los participantes. Al no forzar una asignación estricta de los individuos a un solo *cluster*, se captura mejor la complejidad del movimiento de los jugadores, aumentando la precisión en la clasificación de los patrones de movimiento y minimizando errores que podrían surgir de clasificaciones estrictas.

Por lo tanto, FCM ayudó a identificar patrones de movimiento que pueden no encajar de manera clara en las categorías de jugador o no jugador, lo cual tiene gran valía para detectar los niveles de habilidad progresiva o variaciones en la técnica.

3.4.4 Resultados del análisis categórico

De acuerdo con los resultados de *K-Means* y FCM, se puede asumir lo siguiente:

- 74.2% de los participantes se agrupan con base en la asignación categórica de la variable w_I .
- 2.8% de los participantes no se pudieron categorizar debido a errores en la adquisición del video.
- 22% de los participantes tienen resultados intermedios porque fueron jugadores frecuentes que en la entrevista del protocolo manifestaron estar lesionados.

Se puede observar en estos resultados que alrededor del 75% de los participantes coinciden con la variable w_I , lo que se consideró suficiente para utilizar dicha variable para definir las clases de los videos de la base de datos de DPMov.

IV. FLUJO ÓPTICO PARA EL ANÁLISIS DE M²S

El flujo óptico es un algoritmo útil para el análisis del movimiento de los miembros superiores, ya que permite capturar y cuantificar movimientos detallados durante el proceso de lanzamiento de un balón de Fútbol Americano. Al analizar estos movimientos, se pueden identificar patrones y evaluar la técnica de los lanzamientos. Por lo tanto, este capítulo se centra en el flujo óptico y su aplicación en el análisis del M²S. El algoritmo de Flujo Óptico utilizado es *Gunnar Farnebäck* y se obtienen estadísticas de la media, mediana, entropía y curtosis de las direcciones del movimiento durante el lanzamiento.

El presente capítulo se organiza de la siguiente manera. La sección 4.1 describe el flujo óptico por *Gunnar Farnebäck* y el análisis del M²S. La sección 4.2 presenta las estadísticas obtenidas por el análisis de movimiento. Finalmente, la sección 4.3 presenta la forma de adaptar las estadísticas obtenidas para ser analizadas por modelos de ML.

4.1 Flujo óptico de *Gunnar Farnebäck*

Para capturar y analizar estos movimientos complejos, se utilizó el flujo óptico de *Gunnar Farnebäck*, un método avanzado de visión por computadora. Este método proporciona información detallada sobre el desplazamiento y la dinámica del movimiento, permitiendo extraer estadísticas significativas que pueden ser utilizadas para mejorar el entrenamiento y la prevención de lesiones. Este método se cataloga como flujo óptico denso que estima el desplazamiento de los píxeles entre dos imágenes consecutivas en una secuencia de video. A diferencia de métodos esparcidos que solo estiman el movimiento en algunos puntos, el método de *Gunnar Farnebäck* calcula el flujo óptico en toda la imagen, lo que lo hace muy adecuado para aplicaciones de seguimiento de movimiento en detalle.

El método de *Gunnar Farnebäck* utiliza una aproximación polinómica cuadrática de la variación de intensidad en una pequeña vecindad de píxeles, en lugar de calcular directamente la diferencia entre intensidades. A continuación, se describe su forma de funcionar.

IV. FLUJO ÓPTICO PARA EL ANÁLISIS DE M²S

La imagen en el tiempo t se representa mediante una función de intensidad $I(x, y, t)$ donde x e y son las coordenadas espaciales y t es el tiempo. En cada pequeña vecindad de píxeles, se aproxima la intensidad de la imagen $I(x, y)$ como una superficie cuadrática de la forma:

$$I(x, y) = a_1x^2 + a_2y^2 + a_3xy + a_4x + a_5y + a_6 \quad (4.1)$$

Donde los coeficientes $a_1, a_2, a_3, a_4, a_5, a_6$ se determinan a partir de los datos locales (intensidades de píxeles). Para dos imágenes consecutivas, $I_1(x, y)$ e $I_2(x, y)$, que están separadas por un pequeño intervalo de tiempo, el desplazamiento de los píxeles de una imagen a otra está dado por un vector $d = (dx, dy)$, que indica el movimiento entre las dos imágenes.

Suponiendo que hay una pequeña deformación entre las dos imágenes consecutivas, la relación entre ellas se modela de la siguiente manera:

$$I_2(x, y) \approx I_1(x - dx, y - dy) \quad (4.2)$$

El desplazamiento d entre las dos imágenes se puede obtener ajustando las dos aproximaciones polinómicas (cuadráticas) para minimizar el error en la estimación de intensidades. Es decir, el método busca los vectores de desplazamiento que minimizan la diferencia entre las imágenes aproximadas. Este problema se resuelve de manera eficiente utilizando operaciones de convolución para calcular los coeficientes de las aproximaciones polinómicas y el desplazamiento óptico en cada píxel. Estas convoluciones se definen en [56].

Para hacer el cálculo más robusto y eficiente en diferentes escalas, el método de *Farneback* utiliza un enfoque jerárquico basado en pirámides gaussianas. Se genera una pirámide de imágenes de diferentes resoluciones, comenzando desde una versión más borrosa (imagen reducida) hasta la resolución completa. El flujo óptico se calcula en las imágenes de baja resolución primero, y luego se refina progresivamente en imágenes de mayor resolución. Este enfoque piramidal permite manejar movimientos grandes y pequeños, haciendo el método más robusto frente a grandes desplazamientos. El algoritmo completo viene en [56].

El resultado del método de *Gunnar Farnebäck* es un campo denso de flujo óptico que proporciona un vector de desplazamiento $d = (dx, dy)$ para cada píxel de la imagen. Este vector describe la dirección y magnitud del movimiento local del píxel entre las dos imágenes consecutivas.

Gunnar Farnebäck fue seleccionado para analizar el movimiento porque es particularmente eficaz para analizar movimientos rápidos y complejos, como los que se producen en el lanzamiento de un balón de Fútbol Americano. Además, permite capturar movimientos sutiles con alta precisión y puede manejar variaciones en iluminación y condiciones de captura, lo que es común en ambientes deportivos. Sin embargo, presenta como desventaja que es muy ruidoso.

4.2 Cálculo del flujo óptico

Para el caso de esta tesis, el flujo óptico de *Gunnar Farnebäck* se calculó por medio de los pasos que se describen a continuación.

1. Captura y preprocesamiento de la imagen. Se obtienen dos imágenes consecutivas de una secuencia de video, $I(x, y, t)$ e $I(x, y, t - 1)$, donde x e y son las coordenadas espaciales y t es el tiempo. Estas imágenes se obtienen a partir de buscar el valor máximo del formato RGB de cada píxel dado por:

$$I(x, y, t) = \max [I(x, y, t)^{RGB}] \quad (4.3)$$

2. Suavizado con diferencia de gaussianos (DoG). Las imágenes se suavizan utilizando convolución con un filtro de DoG, que actúa como un detector de bordes y ayuda a reducir el ruido [57]. Esta DoG está dada por:

$$DoG(x, y) = G_{\sigma_1} * I(x, y, t) - G_{\sigma_2} * I(x, y, t - 1) \quad (4.4)$$

Donde G_{σ} es una gaussiana dada por:

$$G_{\sigma} = \exp\left(\frac{x^2 + y^2}{2\sigma^2}\right) \quad (4.5)$$

Los valores de σ se obtienen de [57] y se dejaron como $\sigma_1 = 1$ y $\sigma_2 = 4$. Esta operación ayuda a resaltar las características de bordes y puntos de interés que son útiles para el cálculo de movimiento.

3. Cálculo del flujo óptico de *Gunnar Farnebäck*. Una vez que se obtiene la convolución con la $DoG(x, y)$, se utiliza una aproximación cuadrática local de la intensidad de píxeles para modelar cómo varían las intensidades de una imagen a otra en una pequeña vecindad, la cual se modela como:

$$I(x, y) \approx a_1x^2 + a_2y^2 + a_3xy + a_4x + a_5y + a_6 \quad (4.6)$$

Donde a_1, a_2, a_3 : son coeficientes que modelan la curvatura (componentes cuadráticas) de la intensidad local, a_4, a_5 : son coeficientes que representan los gradientes lineales de la intensidad local y a_6 : representa el valor constante de intensidad en el centro de la vecindad.

En el flujo óptico, se asume que hay una pequeña deformación entre las imágenes consecutivas $I_1(x, y)$ e $I_2(x, y)$. La intensidad de un pixel en una posición (x, y) en la primera imagen cambia ligeramente debido al movimiento. La relación entre las dos imágenes consecutivas es:

$$I_1(x, y) \approx I_2(x - dx, y - dy) \quad (4.7)$$

donde (dx, dy) es el vector de desplazamiento del pixel entre las dos imágenes. Para resolverlo, se utiliza una expansión de Taylor para aproximar la intensidad de $I_2(x - dx, y - dy)$ alrededor de (x, y) :

$$I_2(x - dx, y - dy) \approx I_2(x, y) - \frac{\partial I}{\partial x} dx - \frac{\partial I}{\partial y} dy \quad (4.8)$$

Factorizando la ecuación de intensidad de (4.8), se obtiene:

$$I_1(x, y) - I_2(x, y) \approx \frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy \quad (4.9)$$

Esto genera un sistema de ecuaciones para estimar los desplazamientos dx y dy , que representan el flujo óptico.

Para el caso de esta tesis, se omite la parte de las convoluciones con las gaussianas piramidales por motivos de tiempo de procesamiento. Para omitir estos cálculos sin afectar el rendimiento de *Gunnar Farnebäck*, se utilizó la convolución de los *DoG* definida en (4.4).

En resumen, el cálculo del flujo óptico denso se realiza primero con la convolución de $I(x, y, t)$ e $I(x, y, t - 1)$ con los *DoG* de la ecuación (4.4). Las imágenes suavizadas por las *DoG* se replantean por el modelo cuadrático de la ecuación (4.6) por medio de métodos numéricos que aproximan cada pixel de la imagen a un polinomio de segundo orden. Posteriormente, se realiza el cálculo del desplazamiento óptico por medio de un sistema de ecuaciones que resuelve el modelo de (4.9). Estas ecuaciones se resuelven en cada pixel, produciendo un campo denso de vectores complejos de flujo óptico $d = (dx, dy)$.

4. Cálculo de la magnitud y dirección del flujo óptico. El cálculo de flujo óptico genera un conjunto de números complejos que se deben representar en magnitud y fase. Por lo tanto, se calculan la magnitud y el ángulo de los vectores de desplazamiento obtenidos en cada pixel con la forma:

$$magnitud = \sqrt{dx^2 + dy^2} \quad (4.10)$$

$$dirección = arctang\left(\frac{dy}{dx}\right) \quad (4.11)$$

Estos valores representan la velocidad y la dirección del movimiento para cada píxel en la imagen, respectivamente.

5. Filtrado de ángulos y magnitudes relevantes. El siguiente paso es encontrar las direcciones que describen el movimiento del hombro, brazo, muñeca y mano. Para encontrar estas direcciones se toma el criterio que la lona blanca del protocolo de adquisición del video garantiza que solo existe movimiento derivado del participante. Por lo tanto, si se propone un umbral que elimine ruido y movimientos insignificantes, se pueden encontrar las magnitudes relevantes del movimiento del lanzamiento de pase directo.

De esta manera, se hace un algoritmo simple que revisa las magnitudes del flujo óptico. Si se tienen valores menores o iguales a 0.001 en un determinado pixel, se vuelven cero tanto en

IV. FLUJO ÓPTICO PARA EL ANÁLISIS DE M²S

magnitud como en dirección para dicho pixel. Por otro lado, si las magnitudes son mayores a 0.001, entonces la dirección asociada a ese pixel se mantiene. Las direcciones correspondientes a los píxeles con magnitudes mayores a 0.001 se seleccionan para crear un mapa de vectores de movimiento más relevante. A dichas direcciones se les denominará d en este trabajo.

6. Visualización y análisis. El resultado de las direcciones del flujo óptico d se representa visualmente en un mapa de color que codifica las direcciones en matiz y la magnitud en la saturación de dicho matiz. Se utilizó el círculo cromático (un espectro de matices), donde el matiz de cada píxel representa la dirección del movimiento, mientras que la saturación del color indica la magnitud, tal como se observa en la Figura 4.1.

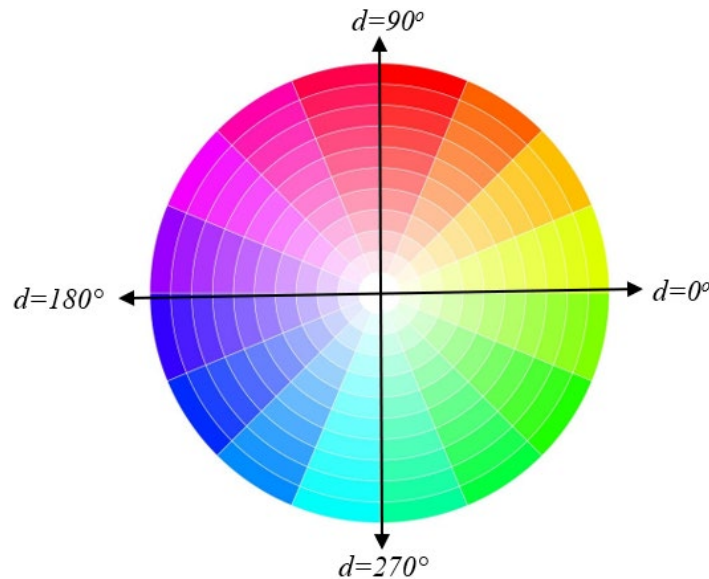


Figura 4.1. Círculo cromático y sus intensidades.

La Figura 4.2 muestra un ejemplo del resultado de aplicar flujo óptico en uno de los videos de la base de datos DPMov. De acuerdo a las imágenes, se visualiza el contorno de los movimientos del individuo en diversos colores. Al compararlo con el círculo cromático, se observa que el color predominante es el anaranjado, lo cual indica que esos vectores de movimiento tienen una dirección aproximada de $d = 65^\circ$. Así se puede estimar la magnitud y dirección de cada movimiento del sujeto.

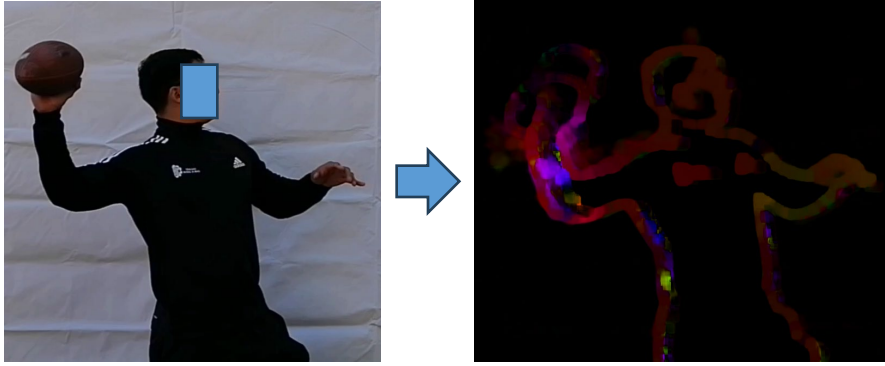


Figura 4.2. En la primera imagen se observa el video regular de un participante, mientras en la segunda imagen se observa su flujo óptico.

4.3 Estadísticas

Las direcciones seleccionadas por magnitudes relevantes obtenidas por el método de *Gunnar Farnebäck* generan un conjunto de datos que se pueden analizar mediante información estadística sin perder información relevante. Por esta razón, en esta sección se revisarán las estadísticas seleccionadas para generar la información que se utilizara como entrada para los modelos de ML.

1. Media del movimiento

La media del movimiento proporciona una medida central de la velocidad y dirección del lanzamiento. Esto ayuda a identificar si los jugadores están utilizando una técnica consistente y eficiente. La media aritmética de un conjunto de datos es la suma de todos los valores dividida por el número total de valores. Matemáticamente, se define como:

$$\mu = \frac{1}{N} \sum_{i=1}^N d_i \quad (4.12)$$

Donde μ es la media del conjunto de datos, N es el número total de observaciones y d_i representa cada uno de los valores de la dirección en un cuadro.

Esta media es proporcional a la velocidad media y la dirección media de los lanzamientos, los cuales describen lo siguiente:

Velocidad media: la media de las velocidades (dy, dx) proporciona información de la consistencia en la fuerza aplicada por el jugador durante el lanzamiento.

Dirección media: la media de las direcciones del movimiento d_i presenta información de cómo se va posicionando el brazo y el cuerpo en cada paso del lanzamiento (el termino paso se refiere a los pasos vistos en la sección 3.2.1).

Si la media muestra poca variabilidad en cada fase del lanzamiento, es un indicio de que el jugador tiene una técnica consistente. Sin embargo, si la media varía significativamente en cada paso del lanzamiento, puede sugerir la necesidad de ajustar o mejorar la técnica.

2. Mediana del movimiento

Es una medida estadística que representa el valor central de un conjunto de datos cuando estos se ordenan de menor a mayor. A diferencia de la media, la mediana no se ve afectada por valores atípicos o extremos, lo que la hace una herramienta robusta para evaluar la distribución de los datos. Esto es útil para evaluar la consistencia del movimiento entre lanzamientos. Matemáticamente, la mediana se define de la siguiente manera:

Para un conjunto de datos con un número impar de observaciones: si N es el número de observaciones, y $d_1, d_2, \dots, d_N$ son las direcciones ordenadas de menor a mayor, entonces la mediana $\tilde{x}$ es el valor que ocupa la posición central:

$$\tilde{x} = d_{(\frac{N+1}{2})} \quad (4.13)$$

Es decir, es el valor en la posición $\frac{N+1}{2}$.

Para un conjunto de datos con un número par de observaciones: la mediana es el valor promedio de los dos valores centrales dado por:

$$\tilde{x} = \frac{d_{(\frac{N}{2})} + d_{(\frac{N}{2}+1)}}{2} \quad (4.14)$$

Cuando se analizan los datos de movimiento obtenidos mediante el método de *Gunnar Farnebäck*, la mediana proporciona una evidencia de la distribución de los desplazamientos, ayudando a evaluar la consistencia del movimiento entre lanzamientos. Es decir, si la mediana de los desplazamientos es constante o monótonica en cada paso del lanzamiento, es porque se tiene una técnica consistente y bien controlada.

Otro aspecto interesante es que, si la media y la mediana son similares, existe una buena técnica de lanzamiento. Por otro lado, conforme vayan difiriendo, significa que la técnica no es adecuada.

3. Entropía del movimiento

La entropía mide el grado de desorden o incertidumbre en el movimiento. Una entropía alta puede indicar variabilidad en la técnica de lanzamiento, lo que podría ser un indicio de falta de entrenamiento o fatiga. La entropía H para un conjunto de datos se define de la siguiente manera:

$$H(d) = - \sum_{i=1}^n p(d_i) \log_2 p(d_i) \quad (4.15)$$

Donde D es una variable aleatoria con n posibles valores (o estados) $d_1, d_2, \dots, d_n$, $p(x_i)$ es la probabilidad de ocurrencia del valor d_i y $\log_2$ es el logaritmo en base 2, que se utiliza para expresar la cantidad de información en bits. De acuerdo con la definición de entropía, se puede considerar lo siguiente:

Entropía baja: sugiere que los movimientos son consistentes y predecibles, lo que sugiere una técnica bien entrenada y una ejecución precisa.

Entropía alta: sugiere que hay variabilidad en la técnica del jugador, posiblemente debido a errores, fatiga, o falta de práctica. Esta variabilidad puede manifestarse como cambios aleatorios en la velocidad o dirección del lanzamiento.

4. Curtosis del movimiento

IV. FLUJO ÓPTICO PARA EL ANÁLISIS DE M²S

La curtosis es una medida estadística que describe la forma de la distribución de un conjunto de datos. En el contexto del análisis de movimiento, la curtosis puede proporcionar información sobre la precisión en la ejecución de los lanzamientos. La curtosis se define como:

$$K = \frac{N \sum_{i=1}^N (d_i - \mu)^4}{(\sum_{i=1}^N (d_i - \mu)^2)^2} - 3 \quad (4.16)$$

Donde N es el número total de observaciones, d_i es la i -ésima dirección, μ es la media de los datos y El término -3 se incluye para que la curtosis de una distribución normal sea 0 (conocido como “exceso de curtosis”). La forma de interpretar la curtosis en este contexto es:

Curtosis baja (platicúrtica): una curtosis baja indica que la distribución tiene una forma plana en comparación con una distribución normal. En el contexto del M²S, una curtosis baja sugiere que existe una mayor dispersión en los desplazamientos musculares durante el lanzamiento. Esto podría ser indicativo de errores en la técnica o de la necesidad de ajustes.

Curtosis alta (leptocúrtica): una curtosis alta indica que los valores de la distribución son cercanos a la media. En el contexto del M²S, una curtosis alta sugiere que la mayoría de los lanzamientos están concentrados alrededor de un valor central. Esto podría indicar una técnica bien definida y consistente.

Este conjunto de medidas estadísticas fue tomado como fundamental en el análisis del rendimiento de los movimientos porque proporcionan un retrato completo de la técnica del jugador. Cada cuadro genera un vector de características estadísticas, por lo que un video genera una matriz de características dado por $F_t = \{\mu_t, \tilde{x}_t, H(d)_t, K_t\}$, donde $t=0, \dots, N-1$ y N es la cantidad de cuadros del video.

La Figura 4.3, muestra un ejemplo de $F_t = \{\mu_t, \tilde{x}_t, H(d)_t, K_t\}$ de un participante de la base de datos DPMov obtenidos a través del flujo óptico. Si se consideran las estadísticas de flujo óptico a lo largo del video, se puede obtener la siguiente información:

- De acuerdo con los pasos del lanzamiento de pase directo, una técnica adecuada consiste en movimiento nulo del brazo en el triángulo Pre-Pase, movimiento a velocidad

constante y con una dirección hacia 90° para la transición a L. Luego, se debe tener movimiento a velocidad constante para la transición a Cero con una dirección que cambia de 90° a 135°. Finalmente, se debe tener una aceleración con velocidad monotónica incremental con dirección que cambia de 135 a 0° durante la extensión. El cuerpo se debe mantener a velocidad constante y con un cambio de dirección en el paso Cero. Esta técnica se refleja en las características de X_t con μ y $\tilde{x}$ aproximadas a la dirección definida en el paso del lanzamiento y con cambios constantes o monotónicos. Por otro lado, $H(d)$ debe ser baja y K leptocúrtica.

- Si la media o la mediana tienen direcciones o velocidades variables en cada paso del lanzamiento, implica una falta de técnica que genera desequilibrio cinético en el brazo que afecta a ligamentos y músculos.
- Si la media y la mediana difieren significa que el brazo y el resto del cuerpo tienen un desfase significativo en velocidad y dirección, causando problemas, principalmente, en el hombro.
- Conforme la entropía aumenta, mayor es la falta de técnica debido a que el brazo y espalda se mueven en una mayor cantidad de direcciones mientras dura el lanzamiento.
- Si la curtosis es platicúrtica, existe un manejo inadecuado en movimientos del miembro superior, lo que genera desgaste en ligamentos, articulaciones y músculos.

4.3 Dimensión de datos

Como se ha mencionado en capítulos anteriores, la base de datos DPMov fue tomada a voluntarias y voluntarios que autorizaron su grabación para los fines de investigación de esta tesis. Debido a que cada persona tenía diferente experiencia con el lanzamiento del balón, cada quien tardaba tiempos diferentes en todo el proceso de lanzamiento, por lo cual, los videos tienen tamaños diferentes. Los videos más cortos tienen 90 cuadros, mientras que los más largos tienen aproximadamente 140 cuadros en total. Esto genera que las matrices X_t sean de tamaño diferente.

IV. FLUJO ÓPTICO PARA EL ANÁLISIS DE M²S

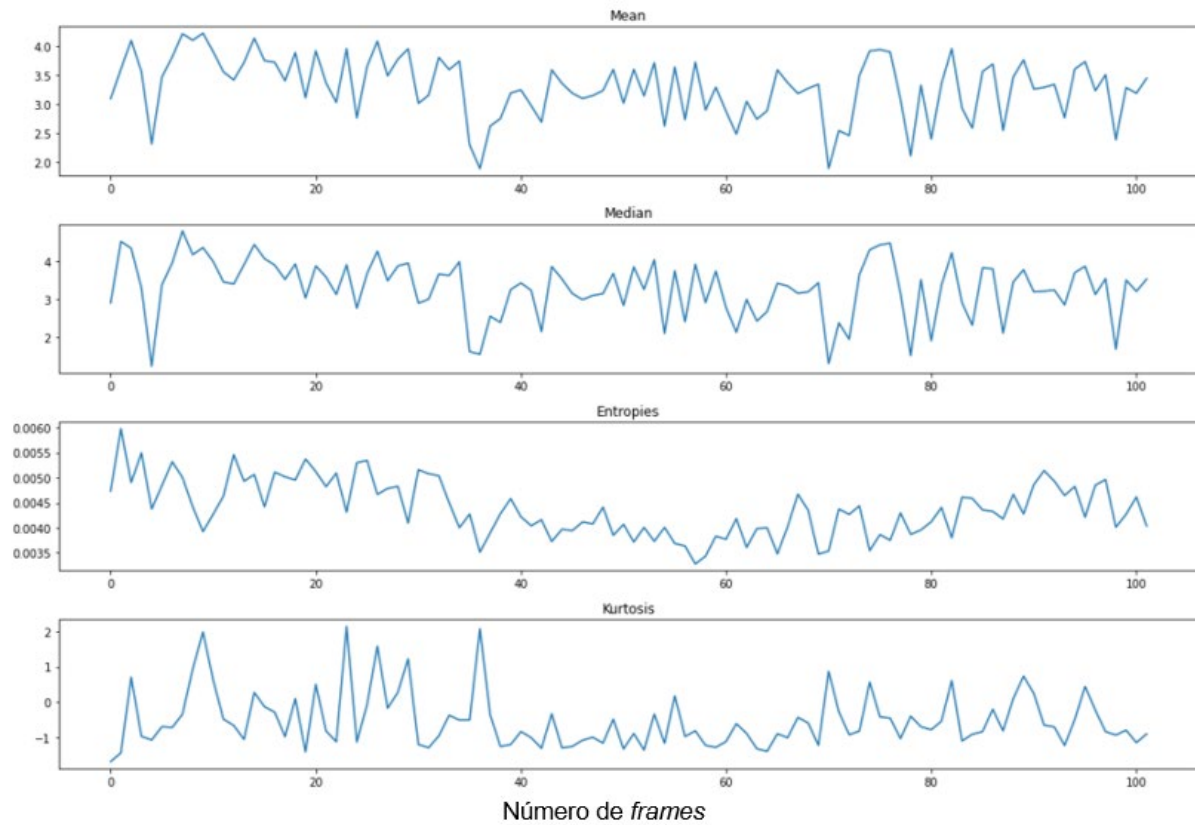


Figura 4.3. Estadísticas de las direcciones obtenidas en el análisis del flujo óptico de un participante de DPmov.

Sin embargo, debido a que los modelos de aprendizaje profundo trabajan con datos de entrada del mismo tamaño, fue necesario redimensionar cada una de las matrices F_t de tamaño $N \times 4$ para dejarlos en un tamaño estándar.

Se optó por utilizar la interpolación lineal para estandarizar F_t a 200 por las siguientes razones:

- Suaviza curvas entre cada cuadro del video y completa datos faltantes.
- Ajusta la longitud de los videos a un tamaño uniforme, garantizando que todos los datos tengan el mismo número de fotogramas.
- Es uno de los métodos de interpolación más simples.
- Se basa en el supuesto de que los puntos dentro del intervalo entre dos datos conocidos pueden ser aproximados por una línea recta.

A continuación, se explica el proceso matemático que se efectúa para conseguir el mismo tamaño para todos los cuadros.

1. Identificación de los marcos originales.

Se entiende que F_t contiene N vectores de características, donde $t = 0, 2, \dots, N - 1$.

2. Definición del nuevo número de datos.

El nuevo número de cuadros deseado se denota como M donde, y para este caso, M es mayor que N . De esta manera, los nuevos índices están uniformemente distribuidos en el intervalo del tiempo original.

3. Cálculo de los nuevos índices.

Se calcula una secuencia de nuevos índices temporales t' que estén espaciados equitativamente entre 0 y $N - 1$, pero con M puntos:

$$t'_i = i \cdot \frac{N - 1}{M - 1}, \quad i = 0, 1, 2, \dots, M - 1 \quad (4.17)$$

Como $M > N$, los nuevos índices incluyen interpolaciones entre los índices originales.

4. Interpolación de los vectores de características.

Para cada nuevo índice t'_i , se encuentra su posición relativa entre dos índices originales consecutivos t_j y t_{j+1} , donde $t_j \leq t'_i \leq t_{j+1}$. El nuevo cuadro $X(t'_i)$ se calcula como una combinación lineal de los fotogramas en t_j y t_{j+1} :

$$X(t'_i) = F(t_j) + \frac{t'_i - t_j}{t_{j+1} - t_j} \cdot (F(t_{j+1}) - F(t_j)) \quad (4.18)$$

Aquí, $F(t_j)$ y $F(t_{j+1})$ son los vectores originales en los tiempos t_j y t_{j+1} .

5. Construcción de la nueva matriz de características.

IV. FLUJO ÓPTICO PARA EL ANÁLISIS DE M²S

Se repite este proceso para todos los M índices para obtener los M vectores ajustados, que constituyen la matriz de características X_t con la longitud deseada.

De esta forma se garantiza que las secuencias de video tengan la misma longitud, debido a que las redes neuronales requieren vectores de entrada de longitud fija.

V. REDES NEURONALES PARA CLASIFICACIÓN DE M²S

En este capítulo se presentará un análisis del desempeño de las redes neuronales artificiales (RNA) para la clasificación del movimiento de miembros superiores (M²S) en jugador y no jugador. La entrada de dichas redes es una matriz de características definida por $X_t = \{x_{1,t}, x_{2,t}, x_{3,t}, x_{4,t}\}$ donde $t = 0, \dots, 199$, $x_1 = \mu$, $x_2 = \tilde{x}$, $x_3 = H(d)$ y $x_4 = K$ y la salida es la variable $w_l = \{\text{jugador}, \text{no jugador}\}$.

A continuación, se expondrán las RNA empleadas en esta tesis, las cuales fueron seleccionadas por ser mayormente apropiadas para el análisis de movimiento de M²S debido a sus capacidades específicas para manejar datos complejos y dinámicos, como los movimientos de los brazos, manos y hombros. Estas redes fueron MLP, CNN y Redes Neuronales Recurrentes (RNN por sus siglas en inglés).

Este capítulo está organizado de la siguiente manera: las secciones 5.1 y 5.2 son un repaso de las redes MLP y CNN mientras que la sección 5.3 aborda brevemente las RNN. Finalmente, la sección 5.4 presenta los resultados de las RNA en la tarea de análisis de M²S.

5.1 Perceptrón multicapa

El MLP es una red neuronal ampliamente utilizada en tareas de aprendizaje supervisado como clasificación y regresión. Es una extensión del perceptrón simple, pero incluye múltiples capas de neuronas organizadas en una arquitectura de red, lo que le permite aprender relaciones complejas y no lineales entre las características de entrada y las salidas esperadas. La estructura del MLP se ejemplifica en la Figura 5.1 y consta de tres tipos de capas:

Capa de entrada. La entrada de la red, que para el caso de esta tesis es la matriz de características X_t .

Capas ocultas. Son las capas intermedias entre la entrada y la salida. Cada capa está compuesta por un grupo de neuronas que generan subespacios que se combinan para formar diversas vertientes de decisión.

Capa de salida. Produce el resultado final, donde las neuronas dependen de las clases de la tarea a solucionar. Para este análisis se utilizan dos neuronas de salida debido a que se tienen dos clases.

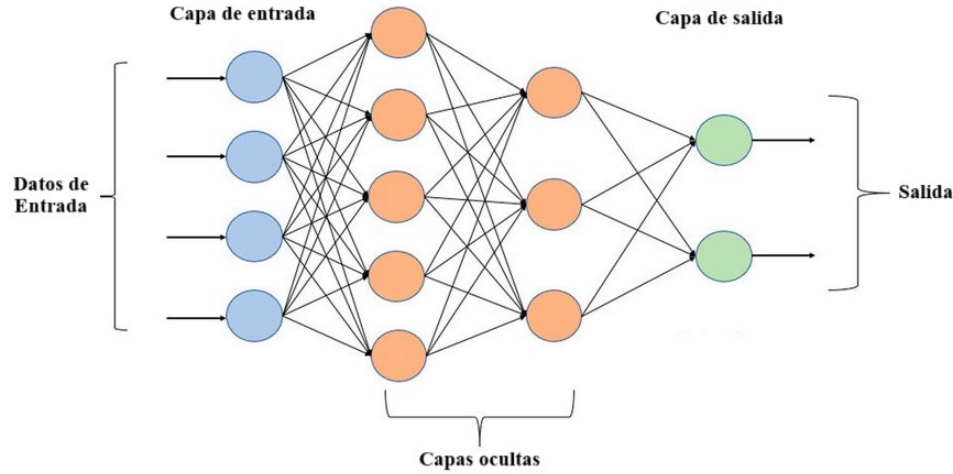


Figura 5.1. Modelo del MLP.

La MLP toma los valores de entrada X_t y cada neurona calcula una combinación lineal con pesos y sesgos w y aplica una función de activación para introducir no linealidad de la siguiente forma:

$$z_{rj} = \sum_{k=0}^P w_{rjk} y_{r-1,jk}, \quad y_{r-1,j0} = +1, \quad y_{rj} = \phi(z_{rj}) \quad (5.1)$$

Donde y_{rj} es j -ésima neurona de la r -ésima capa, $y_{r-1,jk}$ es la capa anterior, donde $y_{0,jk}$ son las entradas X_k , w_k son los pesos asociados a cada entrada y $\phi(z_{rj})$ es la función de activación (sigmoide, ReLu, etc.). En este caso, se utiliza un índice auxiliar k para todas las neuronas y tiene valores de $k=0, 1, \dots, P$.

La forma en cómo funciona la MLP consiste en propagar los datos de entrada hacia adelante (*forward propagation*) para calcular la salida de la red. Los datos de entrada se pasan a través de la primera capa oculta. Si existen más capas ocultas, se calculan las activaciones de las neuronas utilizando las salidas de la capa anterior. Las activaciones de la capa final se interpretan como clasificación o regresión obtenida por la red.

5.1.1 Aprendizaje de la MLP

La MLP es una red supervisada que aprende mediante el algoritmo *backpropagation* que retropropaga el error mediante actualización de pesos definida por:

$$w \leftarrow w - \eta \frac{\partial L}{\partial w} \quad (5.2)$$

Donde η es la tasa de aprendizaje y $\frac{\partial L}{\partial w}$ es el gradiente de la pérdida respecto a w .

El gradiente se actualiza mediante lo siguiente:

1.Función de pérdida. Mide el error de la red. Para el caso de esta tesis, se utilizará la entropía cruzada ya que la capa de salida tiene neuronas de clasificación con función de activación *softmax*.

2.Optimizador. Calcula los gradientes de la función de pérdida con respecto a los pesos y sesgos. Existen diversos optimizadores como ADAM, Nesterov ADAM, Gradiente Descendente Estocástico, propagación RMS, entre otros [58].

3.Propagación. La propagación del error se realiza mediante la regla de la cadena. El gradiente de la pérdida para cualquier peso w_{jr-b} se calcula como:

$$\frac{\partial L}{\partial w_{r-b,j}} = \frac{\partial L}{\partial z_{rj}} \cdot \frac{\partial z_{rj}}{\partial z_{r-1,j}} \cdots \frac{\partial z_{r-b,j}}{\partial w_{r-b,j}} \quad (5.3)$$

Donde $z_{r-b,j}$ es la entrada neta a la neurona j de la capa $r - b$ dada por:

$$z_{r-b,j} = \sum_{k=0}^P w_{r-b,jk} y_{r-b-1,jk} \quad (5.4)$$

$\frac{\partial L}{\partial z_{rj}}$ es el gradiente de la pérdida con respecto a la entrada neta z_{rj} , y se propaga desde la capa de salida hacia atrás.

$\frac{\partial z_{rj}}{\partial z_{r-1,j}}$ se refiere a las propagaciones del error entre capas.

$\frac{\partial z_{r-b,j}}{\partial w_{r-b,j}}$ es la derivada de z_j con respecto a w_{ij} , que equivale a la activación a_i de la neurona de entrada:

$$\frac{\partial z_{r-b,j}}{\partial w_{r-b,j}} = y_{r-b-1,jk} \quad (5.5)$$

5.1.2 Arquitectura de la red

La entrada de la red neuronal contiene los datos estadísticos obtenidos a partir del flujo óptico: media, mediana, entropía y curtosis. Posteriormente, sigue la capa intermedia para la clasificación, lo que conlleva a la capa de salida donde se categoriza en jugador y no jugador.

La entrada de las redes son datos estadísticos de un archivo de Excel que encuentra en cualquiera de dos carpetas: una que engloba los archivos Excel con el X_t de todos los videos de los voluntarios que son “jugador”, así como otra carpeta que abarca los archivos de Excel con las matrices X_t de las personas “no jugador”. Cada Excel contiene las características X_t proporcionadas por el flujo óptico en cada cuadro del video que son: media, mediana, entropía y curtosis. La Tabla 5.1 presenta la forma en cómo se organiza la matriz de características de estadísticas de flujo óptico.

Tabla 5.1. Archivo Excel con las estadísticas de un video del *dataset* DPMov.

	Media	Mediana	Entropía	Curtosis
1	3.0692625	3.1083996	0.00080237	-1.11958822
2	3.0620449	3.1053097	0.00080273	-1.15811079
3	3.0954597	3.1234355	0.00080321	-1.13624661
4	3.0985498	3.1268678	0.00080328	-1.18215284
5	3.1971536	3.2414808	0.0008033	-1.1391373
6	3.1678565	3.1987736	0.00080298	-1.12773282
7	3.0654397	3.1025913	0.00080288	-1.1345409
8	3.0553818	3.0173826	0.00080414	-1.16268211
9	3.1336908	3.1679206	0.00080252	-1.12999828
10	3.166744	3.1818173	0.00080233	-0.99865419

Seguidamente, se tienen dos capas ocultas, una con 192 neuronas y otra con 32 neuronas dadas por la ecuación (5.1). La cantidad de neuronas se seleccionó de manera experimental y

tienen una función ReLU como función de activación dada por la ecuación (2.10). Finalmente, la capa de salida tiene dos neuronas con una *softmax* dada por la ecuación (2.9). A continuación, se presenta un diagrama de la arquitectura de la MLP en la Figura 5.2.

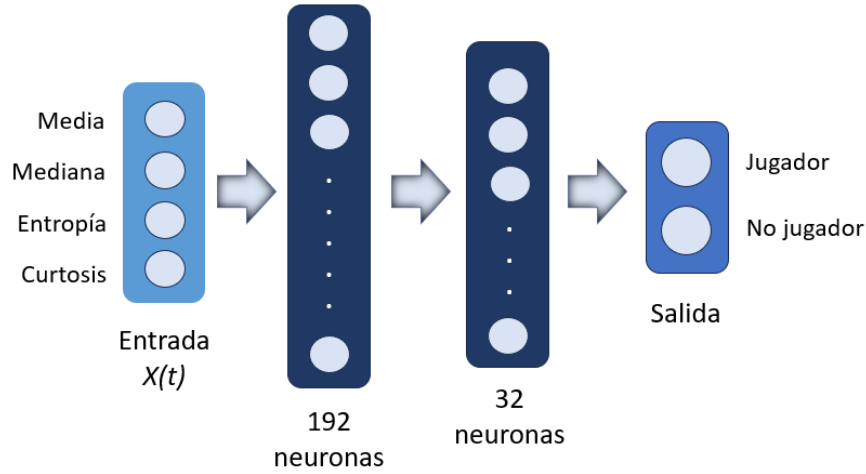


Figura 5.2. Arquitectura de la red neuronal MLP.

5.2 Redes neuronales convolucionales (CNN)

Las CNN son modelos de aprendizaje profundo que contienen un conjunto de capas que pueden ser utilizadas para clasificación y regresión con aprendizaje supervisado basado en *Backpropagation*. Son bastante efectivas para identificar patrones espaciales y características abstractas. Estas redes tienen capas convolucionales para extraer características y capas densamente conectadas para regresión o clasificación. Si bien, en la sección 2.3 de este documento, la CNN para el caso de esta tesis queda planteada de la siguiente manera:

1. Capa convolucional (*convolutional layer*). Realiza operaciones de convolución sobre los datos de la capa anterior para encontrar características abstractas. Para el caso del análisis de M²S, estas capas están dadas por:

$$y_{rj} = f \left[\sum_{p=0}^{k-1} \sum_{q=0}^{k-1} Y_{r-1,j,m-p,n-q} W_{r,j,m,n} + b_{r,j} \right] \quad (5.6)$$

Donde p y q son variables auxiliares para modelar la convolución, $Y_{r-1,j}$ es la salida de la capa anterior, $W_{r,j,m,n}$ es el kernel de la j -ésima neurona de la r -ésima capa y $b_{r,j}$ son los sesgos. Para la primera capa, $Y_{r-1,j}$ es X_t . La función de activación es ReLU dada por la ecuación (2.10).

2.Capa de *pooling* (*pooling layer*). Reduce la dimensionalidad de los mapas de características mediante operaciones de *max-pooling* dado por la ecuación (2.6).

3.*Dropout*. Es una capa que apaga aleatoriamente neuronas durante el entrenamiento para reducir el sobreajuste. La entrada del *Dropout* es el porcentaje de neuronas a apagar.

4.Capa completamente conectada (*fully connected layer*). Los mapas de características son aplanados en un vector y pasados a capas completamente conectadas dados por la ecuación (5.1). La función de activación de las capas totalmente conectadas ocultas es un ReLU. No obstante, la última capa tiene *softmax* dada por la ecuación (2.9) como función de activación.

5.2.1 Arquitectura

Para esta tesis, se desarrollaron dos arquitecturas de redes CNN que se llamaron CNN+MLP y DCNN.

CNN+MLP es una red que tiene como entrada la matriz de características X_t . La parte de extracción de características es una capa convolucional que contiene 16 neuronas compuestas de kernels $W_{1,j}$ de 3x3, función de activación ReLU y con una capa de *max-pooling* de 2x2. La capa de clasificación tiene la misma estructura de la MLP de la sección 5.1.1.

DCNN es una red profunda con varias capas de convolucionales en la parte de extracción de características. Su entrada es una matriz de características X_t . La parte de extracción de características tiene la siguiente arquitectura:

- Capa de 8 neuronas compuestas de kernels $W_{1,j}$ de 3x3, función de activación ReLU y con una capa de *max-pooling* de 2x2 y un *Dropout* de 20%.
- Capa de 16 neuronas compuestas de kernels $W_{2,j}$ de 3x3, función de activación ReLU y con una capa de *max-pooling* de 2x2 y un *Dropout* de 20%.

- Capa de 64 neuronas compuestas de kernels $W_{3,j}$ de 3x3, función de activación ReLU y con una capa de *max-pooling* de 2x2 y un *Dropout* de 20%.

La capa de clasificación tiene la misma estructura de la MLP de la sección 5.1.1.

5.3 Redes neuronales recurrentes

Las redes neuronales recurrentes (RNN por sus siglas en inglés) son modelos que conectan de forma organizada unidades llamadas neuronas recurrentes. Una neurona recurrente actualiza su estado con base en la entrada actual y su estado anterior, tal como se ve en la Figura 5.3. El cálculo de estas neuronas se realiza en forma iterativa y de la siguiente manera:

$$y_t = f(w_{hb}y_{t-b} \dots + w_{h1}y_{t-1} + w_hx_t + w_0) \quad (5.7)$$

Donde w_h son los pesos recurrentes conectados con el estado anterior de la neurona, w_x son los pesos de entrada y w_o el sesgo de la neurona.

Las RNN tienen bucles internos que permiten tener un estado interno de memoria y se utilizan para procesar datos secuenciales, series de tiempo o información donde importa el orden en que aparecen los datos. Los tipos más comunes son las RNN simples y las memorias de corto y largo plazo (LSTM por sus siglas en inglés), las cuales se describirán brevemente.

5.3.1 RNN simple

Una RNN simple es la versión más sencilla de una RNN. En esta arquitectura, cada neurona tiene una conexión recurrente que le permite mantener un estado y_t , actualizarlo en cada paso de tiempo en función de la entrada actual x_t y el estado previo y_t , tal como se observa en la siguiente expresión:

$$y_t = f(w_{h1}y_{t-1} + w_hx_t + w_0) \quad (5.16)$$

Estas neuronas se pueden conectar de forma recurrente con otras neuronas como se observa en la Figura 5.3.

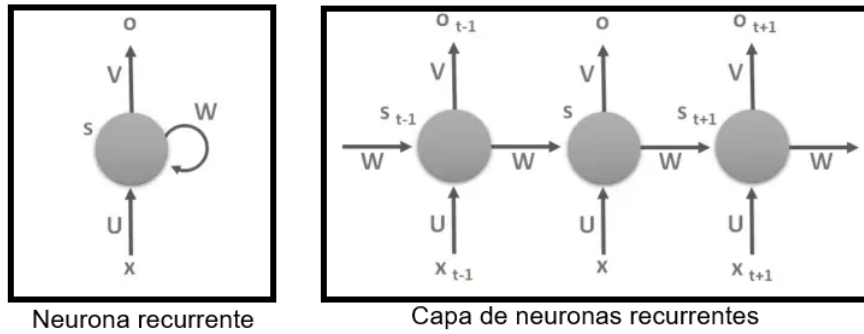


Figura 5.3. Redes neuronales recurrentes simples.

Para el caso de esta tesis se implementó una RNN simple con la siguiente estructura:

- Su entrada es una matriz de características X_t .
- Capa de RNN simple con 128 neuronas y función de activación ReLU para extraer características.
- Capa de RNN simple con 64 neuronas y función de activación ReLU para extraer características.
- La capa de clasificación tiene la misma estructura de la MLP de la sección 5.1.1.

5.3.2 Memorias de corto y largo plazo

La arquitectura de una LSTM introduce un estado de memoria (c_t) y mecanismos llamados puertas que controlan la información que fluye por la red. Esto le permite a cada neurona tener control de cuando aprender, cuándo olvidar, cuándo recordar y cuándo actualizar la información. Los componentes de la LSTM son el estado y_t y la puerta de olvido, de entrada y de salida [59]. En la Figura 5.4 se puede observar su estructura.

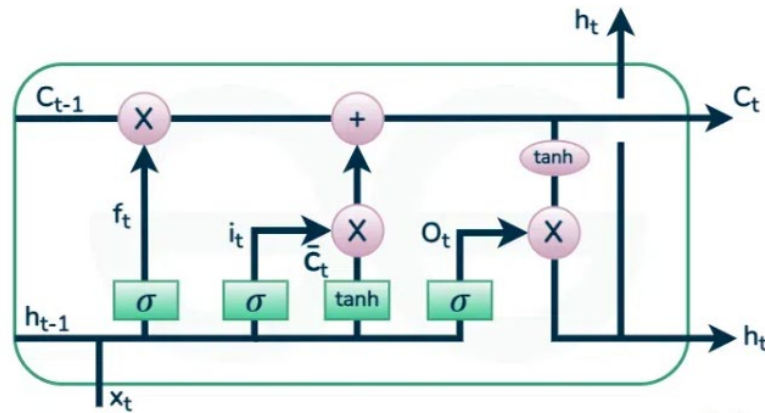


Figura 5.4. Unidad de una LSTM.

Para el caso de esta tesis se implementó una RNN simple con la siguiente estructura:

- Su entrada es una matriz de características X_t .
- Capa de 128 neuronas LSTM con la estructura de la Figura 5.4.
- La capa de clasificación tiene la misma estructura de la MLP de la sección 5.1.1.

5.4 Resultados

5.4.1 Parámetros de aprendizaje

Todas las redes se entrenaron con el optimizador ADAM debido a su rápida convergencia y ajuste dinámico para cada parámetro del modelo. Como función de pérdida, se utilizó la entropía cruzada (*cross-entropy*) ya que mide de manera directa la discrepancia entre las distribuciones de probabilidad predichas por el modelo y las distribuciones verdaderas. El *batch* se fijó en 8. Las épocas varían según el modelo.

Las métricas utilizadas en esta sección de resultados son exactitud (Acc), medición F (F1), precisión (P) y sensibilidad (R) definidas en [58]. Estas métricas fueron seleccionadas porque son las más populares en clasificación binaria.

5.4.2 Resultados de las métricas

Las redes que se emplearon fueron MLP, CNN+MLP, DCNN, LSTM y RNN simple. La Tabla 5.1 muestra los resultados de las métricas anteriormente indicadas.

V. REDES NEURONALES PARA CLASIFICACIÓN DE M²S

Tabla 5.1. Métricas de las RNA.

Red	Acc	F1	P	R
MLP	0.75	0.64	0.78	0.54
CNN+MLP	0.75	0.60	0.86	0.46
DCNN	0.75	0.64	0.78	0.54
LSTM	0.81	0.77	0.77	0.77
RNN Simple	0.62	0.67	0.71	0.63

En la comparación de los resultados, se puede observar que la RNA que proporciona mejores resultados es LSTM debido a su exactitud y medición, además de que la precisión y sensibilidad consiguen una proporción equivalente. En cuanto a Acc y F1 le prosiguen la MLP y DCNN. En tercer lugar, se encuentra CNN+MLP, sin embargo, P y R son muy diferentes, lo que genera incertidumbre en saber si realmente clasifica. Finalmente, la RNN simple es la más baja, pero su P y R están balanceados, lo que quiere que tiene aprendizaje correcto.

VI. CONCLUSIONES

En esta tesis se presentó una metodología de ciencia de datos para el análisis del movimiento de miembros superiores durante el pase directo de balón de Fútbol Americano. La metodología gestionada está constituida por diversos aspectos: base de datos, análisis categórico, estadísticas, flujo óptico, y redes neuronales artificiales.

La base de datos se llamó DPMov y está conformada por 106 videos obtenidos de 35 participantes. Los videos fueron tomados utilizando un protocolo que consta de los siguientes elementos para la toma de video: fondo blanco, cámara con resolución de 1080 pixeles, trípode para mantener un enfoque fijo y toma de datos de los participantes. Los datos requeridos por los participantes fueron edad, peso, estatura, sexo, brazo dominante, lesiones y si eran jugadores habituales o no.

Una vez que se generó la base de datos, se realizó un análisis multivariado para encontrar un vector de características que conlleve a un análisis categórico basado en agrupamiento de *K-Means* y FCM. A partir del análisis categórico, se decidió utilizar la variable de $wI = \{\text{jugador/no jugador}\}$ para definir las clases contenidas en DPMov. Para analizar el M2S a partir de los videos, se utilizó el algoritmo de Gunnar Farnebäck, el cual permite encontrar las direcciones relevantes al movimiento de miembros superiores. El siguiente paso es encontrar una matriz de características F_t con la media, mediana, entropía y curtosis del flujo óptico. Esta matriz se redimensiona a 200x4 y se le denomina X_t , X_t es la entrada de las RNA. Las RNA que se eligieron en esta investigación tienen arquitecturas basadas en MLP, CNN y RNN.

De acuerdo con los resultados obtenidos en el análisis de las RNA, la red LSTM fue la que mejor obtuvo resultados como se puede observar en Acc y F1, además, se puede observar por la P y la R que la red está aprendiendo. En última instancia se tienen las RNN que, no obstante, su exactitud y medición F son las menores en comparación con las demás redes, demuestran que su aprendizaje es mayor que DCNN y MLP, aunque sus mediciones en exactitud y medición F sean mayores. CNN+MLP tiene un Acc y F1 bajos y existe evidencia de su bajo aprendizaje. Por lo tanto, se sugiere que las RNN generan evidencia de que aprenden la tarea de M²S con el vector X_t . También se sugiere que LSTM es la opción más adecuada debido a

VI. CONCLUSIONES

que proporciona una herramienta robusta y eficaz para modelar las dependencias temporales complejas del movimiento, superando las limitaciones que presentan otras arquitecturas.

Los diversos experimentos y la información recabada en esta tesis respecto a los participantes y el análisis categórico mostraron correlaciones que generaron que sugieren lo siguiente:

- De acuerdo con los pasos del lanzamiento de pase directo, una técnica adecuada consiste en mantener un cuerpo con movimientos a velocidad constante y un cambio de dirección en el paso Cero. El brazo debe tener velocidad nula, constante y monotónica incremental con cambios de dirección constantes. Esta técnica se refleja en las características de X_t con μ y $\tilde{x}$ constantes o monotónicas con direcciones definidas por los pasos del lanzamiento, $H(d)$ baja y K leptocúrtica.
- Si la media o la mediana son variables en cada paso del lanzamiento, implica que la técnica no es adecuada debido a variaciones en velocidad de lanzamiento, lo que afecta ligamentos principalmente del hombro.
- Si la media y la mediana difieren significa que la posición del cuerpo y el brazo son divergentes, por lo que existe impacto en todas las partes implicadas de miembros superiores.
- Si la entropía es alta, se observa falta de práctica lo que implica que existen movimientos erráticos, los cuales afectan en ligamentos y músculos articulares.
- Si la curtosis es platicúrtica existe una mayor dispersión en los movimientos musculares por lo que se puede observar que en cada parte de los miembros musculares existe alguna disociación.

De acuerdo con los resultados obtenidos, se considera que esta metodología de ciencia de datos generó un prototipo de análisis de movimiento de miembros superiores en nivel de madurez tecnológica (TRL, por sus siglas en inglés *Technological Readiness Level*) de TRL-4 en las escalas de NASA, Unión Europea y Departamento de Energía de Estados Unidos. El criterio para categorizar el prototipo en este nivel se debe a que la base de datos y la integración de los componentes del prototipo se realizaron en entornos de laboratorio, condiciones

VI. CONCLUSIONES

controladas y se obtuvieron los resultados esperados. Por lo tanto, el trabajo a futuro es desarrollar el prototipo hasta el nivel TR-6 que considera resultados de validación adecuada en entornos reales. Para ello, se debe realizar una base de datos en entornos de prácticas reales de Fútbol Americano con un entrenador certificado y desarrollar una red neuronal recurrente basada en neuronas LSTM para el análisis de cada una de las etapas del movimiento durante el pase directo.

VII. REFERENCIAS

- [1] Fang, H., Li, J., Tang, H., Xu, C., Zhu, H., Xiu, Y., Li, Y., & Lu, C. (2022). AlphaPose: Whole-Body Regional Multi-Person Pose Estimation and Tracking in Real-Time. *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 45(6), 7157-7173. <https://doi.org/10.1109/tpami.2022.3222784>
- [2] Singh, A., Bevilacqua, A., Nguyen, T. L., Hu, F., McGuinness, K., O'Reilly, M., Whelan, D., Caulfield, B., & Ifrim, G. (2022). Fast and robust video-based exercise classification via body pose tracking and scalable multivariate time series classifiers. *Data Mining And Knowledge Discovery*, 37(2), 873-912. <https://doi.org/10.1007/s10618-022-00895-4>
- [3] Zhang, J., Li, P., Jin, C., Zhang, W., & Liu, S. (2019). A Novel Adaptive Kalman Filtering Approach to Human Motion Tracking With Magnetic-Inertial Sensors. *IEEE Transactions On Industrial Electronics*, 67(10), 8659-8669. <https://doi.org/10.1109/tie.2019.2946557>
- [4] Digo, E., Gastaldi, L., Antonelli, M., Pastorelli, S., Cereatti, A., & Caruso, M. (2022). Real-time estimation of upper limbs kinematics with IMUs during typical industrial gestures. *Procedia Computer Science*, 200, 1041-1047. <https://doi.org/10.1016/j.procs.2022.01.303>
- [5] Yang, Y., Ren, Z., Li, H., Zhou, C., Wang, X., & Hua, G. (2021). Learning Dynamics via Graph Neural Networks for Human Pose Estimation and Tracking. *2022 IEEE/CVF Conference On Computer Vision And Pattern Recognition (CVPR)*. <https://doi.org/10.1109/cvpr46437.2021.00798>
- [6] Velarde, M., Perugachi, E., Romero, D., Sappa, A., & Vintimilla, B. (2015). Análisis del movimiento de las extremidades superiores aplicado a la rehabilitación física de una persona usando técnicas de visión artificial. *Revista Tecnológica*, 28.
- [7] R. L. A. Michel Latarjet, «Miembro Superior,» de *Anatomía Humana*, Medica Panamericana, 2019.
- [8] R. H. González, Artist, *Apuntes Generales de Anatomía Morfológica Aplicada Osteología*. [Art]. Universidad Complutense de Madrid, 2012.
- [9] V. Silberman, *Ortopedia Y Traumatología*, Editorial Médica Panamericana, 2017.
- [10] Kelly, B. T., Barnes, R. P., Powell, J. W., & Warren, R. F. (2004). Shoulder Injuries to Quarterbacks in the National Football League. *The American Journal Of Sports Medicine*, 32(2), 328-331. <https://doi.org/10.1177/0363546503261737>

VI. REFERENCIAS

- [11] Martínez, A. M., Espinosa, J. G., Santiago, F. R., Del Mar Castellano García, M., & Álvarez, L. G. (2018). Hallazgos por imagen de las lesiones deportivas del miembro superior. *Seram*. <https://www.piper.espacio-seram.com/index.php/seram/article/download/84/83>
- [12] Hoppe, M.W., Brochhagen, J., Tischer, T. *et al.* Risk factors and prevention strategies for shoulder injuries in overhead sports: an updated systematic review. *J EXP ORTOP* 9, 78 (2022).
- [13] Confino, J., Irvine, J. N., O'Connor, M., Ahmad, C. S., & Lynch, T. S. (2019). Early Sports Specialization Is Associated With Upper Extremity Injuries in Throwers and Fewer Games Played in Major League Baseball. *Orthopaedic Journal Of Sports Medicine*, 7(7). <https://doi.org/10.1177/2325967119861101>
- [14] Giannatos, V., Panagopoulos, A., Antzoulas, P., Giakoumakis, S. I., Lakoumentas, J., & Kouzelis, A. (2022). Functional performance of the upper limb and the most common boxing-related injuries in male boxers: a retrospective, observational, comparative study with non-boxing population. *BMC Sports Science Medicine And Rehabilitation*, 14(1). <https://doi.org/10.1186/s13102-022-00558-3>
- [15] G. L. S. de las Heras Paz María Mercedes, «Lesiones Comunes en el Aparato Musculoesquelético y Estrategias de Prevención,» *Ocronos*, vol. 6, nº 12, 2023.
- [16] A. C. Bovik, *The Essential Guide to Video Processing*, USA: Academic Press, Inc., 2009.
- [17] M. Leo, G. Medioni, M. Trivedi, T. Kanade, & G.M. Farinella. (2017). Computer vision for assistive technologies. *Computer Vision And Image Understanding*, 154, 1-15.
- [18] Lucas, B. D. (1985). *Generalized image matching by the method of differences* [Ph.D. dissertation, Carnegie-Mellon University]. <https://ci.nii.ac.jp/naid/10011856486>
- [19] Lucas, Bruce D., & Kanade, Takeo. (1981). *An iterative image registration technique with an application to stereo vision*. Proceedings Of The 7th International Joint Conference On Artificial Intelligence - Volume 2.
- [20] G. Farneback, «Polynomial Expansion for Orientation and Motion Estimation,» 2002.
- [21] G. Farneback, «Two-Frame Motion Estimation Based on Polynomial Expansion,» de *Image Analysis*, Berlin, Heidelberg, Springer Berlin Heidelberg, 2003, pp. 363-370.
- [22] «Inteligencia Artificial,» México, Prentice-Hall, 2016, pp. 113-141.

VI. REFERENCIAS

- [23] Rocklöv, J., & Gayle, A. A. (2020). Data Science and Machine Learning: Mathematical and Statistical Methods. *International Journal Of Epidemiology*, 49(6), 2096. <https://doi.org/10.1093/ije/dyaa111>
- [24] M. Brodie, «What Is Data Science?,» 2019, pp. 101-130.
- [25] A. C. Franco, «Deep Learning,» Sevilla, 2019.
- [26] G. Bonaccorso, Machine Learning Algorithms, Birmingham: Packt Publishing Ltd, 2017.
- [27] Y. B. A. C. Ian Goodfellow, Deep Learning, MIT Press, 2016.
- [28] S. Haykin, Neural Networks. A Comprehensive Foundation, Ontario, Canada: Pearson Education, 1999.
- [29] Badillo, F. L., Hernández, C. A. R., Narváez, B. M., & Trillos, Y. E. A. (2021). Redes neuronales convolucionales: un modelo de Deep Learning en imágenes diagnósticas. Revisión de tema. *Revista Colombiana de Radiología*, 32(3), 5591-5599. <https://doi.org/10.53903/01212095.161>
- [30] Yasaka, K., Akai, H., Kunimatsu, A., Kiryu, S., & Abe, O. (2018). Deep learning with convolutional neural network in radiology. *Japanese journal of radiology*, 36(4), 257–272. <https://doi.org/10.1007/s11604-018-0726-3>
- [31] Philp, F., Freeman, R., & Stewart, C. (2022). An international survey mapping practice and barriers for upper-limb assessments in movement analysis. *Gait & Posture*, 96, 93-101. <https://doi.org/10.1016/j.gaitpost.2022.05.018>
- [32] Palermo, M., Cerqueira, S., André, J., Pereira, A., & Santos, C. (2022). Complete Inertial Pose Dataset: from raw measurements to pose with low-cost and high-end MARG sensors. *ArXiv*. <https://doi.org/10.48550/arXiv.2202.06164>
- [33] Li, R., Jumet, B., Ren, H., Song, W., & Tse, Z. T. H. (2020). An inertial measurement unit tracking system for body movement in comparison with optical tracking. *Proceedings Of The Institution Of Mechanical Engineers Part H Journal Of Engineering In Medicine*, 234(7), 728-737. <https://doi.org/10.1177/0954411920921695>
- [34] Mira, R. M., Tosatti, L. M., Sacco, M., & Scano, A. (2021). Detailed characterization of physiological EMG activations and directional tuning of upper-limb and trunk muscles in point-to-point reaching movements. *Current Research In Physiology*, 4, 60-72. <https://doi.org/10.1016/j.crphys.2021.02.005>

VI. REFERENCIAS

- [35] Islam, M. R. U., & Bai, S. (2022). A novel approach of FMG sensors distribution leading to subject independent approach for effective and efficient detection of forearm dynamic movements. *Biomedical Engineering Advances*, 4, 100062. <https://doi.org/10.1016/j.bea.2022.100062>
- [36] Ohashi, T., Ikegami, Y., & Nakamura, Y. (2020). Synergetic reconstruction from 2D pose and 3D motion for wide-space multi-person video motion capture in the wild. *Image And Vision Computing*, 104, 104028. <https://doi.org/10.1016/j.imavis.2020.104028>
- [37] Wang, M., Tighe, J., & Modolo, D. (2020). Combining Detection and Tracking for Human Pose Estimation in Videos. *2022 IEEE/CVF Conference On Computer Vision And Pattern Recognition (CVPR)*. <https://doi.org/10.1109/cvpr42600.2020.01110>
- [38] Mohseni, M., Shalchyan, V., Jochumsen, M., & Niazi, I. K. (2019). Upper limb complex movements decoding from pre-movement EEG signals using wavelet common spatial patterns. *Computer Methods And Programs In Biomedicine*, 183, 105076. <https://doi.org/10.1016/j.cmpb.2019.105076>
- [39] Kim, J., Li, G., Yun, I., Jung, C., & Kim, J. (2021). Weakly-supervised temporal attention 3D network for human action recognition. *Pattern Recognition*, 119, 108068. <https://doi.org/10.1016/j.patcog.2021.108068>
- [40] Ji, X., Zhao, Q., Cheng, J., & Ma, C. (2021). Exploiting spatio-temporal representation for 3D human action recognition from depth map sequences. *Knowledge-Based Systems*, 227, 107040. <https://doi.org/10.1016/j.knosys.2021.107040>
- [41] Xu, J., Song, R., Wei, H., Guo, J., Zhou, Y., & Huang, X. (2020). A fast human action recognition network based on spatio-temporal features. *Neurocomputing*, 441, 350-358. <https://doi.org/10.1016/j.neucom.2020.04.150>
- [42] Xiong, Q., Zhang, J., Wang, P., Liu, D., & Gao, R. X. (2020). Transferable two-stream convolutional neural network for human action recognition. *Journal Of Manufacturing Systems*, 56, 605-614. <https://doi.org/10.1016/j.jmsy.2020.04.007>
- [43] Ye, Q., Tan, Z., & Zhang, Y. (2022). Human action recognition method based on Motion Excitation and Temporal Aggregation module. *Heliyon*, 8(11), e11401. <https://doi.org/10.1016/j.heliyon.2022.e11401>
- [44] Zhang, Y., Guo, Q., Du, Z., & Wu, A. (2023). Human Action Recognition for Dynamic Scenes of Emergency Rescue Based on Spatial-Temporal Fusion Network. *Electronics*, 12(3), 538. <https://doi.org/10.3390/electronics12030538>

VI. REFERENCIAS

- [45] Beddiar, D. R., Oussalah, M., & Nini, B. (2021). Fall detection using body geometry and human pose estimation in video sequences. *Journal Of Visual Communication And Image Representation*, 82, 103407. <https://doi.org/10.1016/j.jvcir.2021.103407>
- [46] Host, K., & Ivašić-Kos, M. (2021). An overview of Human Action Recognition in sports based on Computer Vision. *Heliyon*, 8(6). <https://doi.org/10.1016/j.heliyon.2022.e09633>
- [47] Zhou, Y., Chen, C., Cheng, M., Alshahrani, Y., Franovic, S., Lau, E., Xu, G., Ni, G., Cavanaugh, J. M., Muh, S., & Lemos, S. (2021). Comparison of machine learning methods in sEMG signal processing for shoulder motion recognition. *Biomedical Signal Processing And Control*, 68, 102577. <https://doi.org/10.1016/j.bspc.2021.102577>
- [48] Podoprosvetov, A., Alisejchik, A., & Orlov, I. (2021). Comparison of action recognition from video and IMUs. *Procedia Computer Science*, 186, 242-249. <https://doi.org/10.1016/j.procs.2021.04.144>
- [49] Scano, A., Dardari, L., Molteni, F., Giberti, H., Tosatti, L. M., & D'Avella, A. (2019). A Comprehensive Spatial Mapping of Muscle Synergies in Highly Variable Upper-Limb Movements of Healthy Subjects. *Frontiers In Physiology*, 10. <https://doi.org/10.3389/fphys.2019.01231>
- [50] Engdahl, S. M., & Gates, D. H. (2019). Reliability of upper limb movement quality metrics during everyday tasks. *Gait & Posture*, 71, 253-260. <https://doi.org/10.1016/j.gaitpost.2019.04.023>
- [51] Xiaobo Zhou, Renling Zou, & Xiayang Huang. (2021). Single upper limb functional movements decoding from motor imagery EEG signals using wavelet neural network. *Biomedical Signal Processing And Control*, 70. <https://doi.org/10.1016/j.bspc.2021.102965>
- [52] Dorschky, E., Camomilla, V., Davis, J., Federolf, P., Reenalda, J., & Koelewijn, A. D. (2022). Perspective on “in the wild” movement analysis using machine learning. *Human Movement Science*, 87, 103042. <https://doi.org/10.1016/j.humov.2022.103042>
- [53] Cronin, N. J. (2021). Using deep neural networks for kinematic analysis: Challenges and opportunities. *Journal Of Biomechanics*, 123, 110460. <https://doi.org/10.1016/j.jbiomech.2021.110460>
- [54] D. Maddox, *Adapt or Die: Advancements to Accelerate Execution in Football's Modern Passing Game*, Certa Publishing, 2018.

VI. REFERENCIAS

- [55] Kumar, P., & Kanavalli, A. (2021). A Similarity based K-Means Clustering Technique for Categorical Data in Data Mining Application. *International Journal Of Intelligent Engineering And Systems*, 14(2), 43-51. <https://doi.org/10.22266/ijies2021.0430.05>
- [56] Farneback, G. (2002). Very high accuracy velocity estimation using orientation tensors, parametric motion, and simultaneous segmentation of the motion field. *Proceedings Of The IEEE International Conference On Computer Vision*, 1, 171-177. <https://doi.org/10.1109/iccv.2001.937514>
- [57] J. A. R. Quintana, «Diseño de redes neuronales basados en modelos perceptuales multinivel para objetos estáticos y dinámicos en secuencias de video,» Disertación doctoral, Tecnológico Nacional de México, Chihuahua, México, 2014.
- [58] I. a. B. Y. a. C. A. Goodfellow, Deep Learning, MIT Press, 2016.
- [59] Ghojogh, B., & Ghodsi, A. (2023). Recurrent Neural Networks and Long Short-Term Memory Networks: Tutorial and Survey. *ArXiv*. <https://doi.org/10.48550/arXiv.2304.11461>
- [60] Li, R., Jumet, B., Ren, H., Song, W., & Tse, Z. T. H. (2020b). An inertial measurement unit tracking system for body movement in comparison with optical tracking. *Proceedings Of The Institution Of Mechanical Engineers Part H Journal Of Engineering In Medicine*, 234(7), 728-737. <https://doi.org/10.1177/0954411920921695>
- [61] Zhang, J., Li, P., Jin, C., Zhang, W., & Liu, S. (2019b). A Novel Adaptive Kalman Filtering Approach to Human Motion Tracking With Magnetic-Inertial Sensors. *IEEE Transactions On Industrial Electronics*, 67(10), 8659-8669. <https://doi.org/10.1109/tie.2019.2946557>
- [62] Wang, M., Tighe, J., & Modolo, D. (2020d). Combining Detection and Tracking for Human Pose Estimation in Videos. *2022 IEEE/CVF Conference On Computer Vision And Pattern Recognition (CVPR)*. <https://doi.org/10.1109/cvpr42600.2020.01110>
- [63] Palermo, M., Cerqueira, S., De André, J. P. A. M., Pereira, A., & Peixoto Dos Santos, C. (2022). Complete Inertial Pose Dataset: from raw measurements to pose with low-cost and high-end MARG sensors. *ArXiv*.
- [64] Mira, R. M., Tosatti, L. M., Sacco, M., & Scano, A. (2021b). Detailed characterization of physiological EMG activations and directional tuning of upper-limb and trunk muscles in point-to-point reaching movements. *Current Research In Physiology*, 4, 60-72. <https://doi.org/10.1016/j.crphys.2021.02.005>

VI. REFERENCIAS

- [65] Fang, H., Li, J., Tang, H., Xu, C., Zhu, H., Xiu, Y., Li, Y., & Lu, C. (2022b). AlphaPose: Whole-Body Regional Multi-Person Pose Estimation and Tracking in Real-Time. *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 45(6), 7157-7173. <https://doi.org/10.1109/tpami.2022.3222784>
- [66] Yang, Y., Ren, Z., Li, H., Zhou, C., Wang, X., & Hua, G. (2021b). Learning Dynamics via Graph Neural Networks for Human Pose Estimation and Tracking. *2022 IEEE/CVF Conference On Computer Vision And Pattern Recognition (CVPR)*. <https://doi.org/10.1109/cvpr46437.2021.00798>
- [67] Islam, M. R. U., & Bai, S. (2022b). A novel approach of FMG sensors distribution leading to subject independent approach for effective and efficient detection of forearm dynamic movements. *Biomedical Engineering Advances*, 4, 100062. <https://doi.org/10.1016/j.bea.2022.100062>
- [68] Digo, E., Gastaldi, L., Antonelli, M., Pastorelli, S., Cereatti, A., & Caruso, M. (2022b). Real-time estimation of upper limbs kinematics with IMUs during typical industrial gestures. *Procedia Computer Science*, 200, 1041-1047. <https://doi.org/10.1016/j.procs.2022.01.303>
- [69] Zhou, Y., Chen, C., Cheng, M., Alshahrani, Y., Franovic, S., Lau, E., Xu, G., Ni, G., Cavanaugh, J. M., Muh, S., & Lemos, S. (2021b). Comparison of machine learning methods in sEMG signal processing for shoulder motion recognition. *Biomedical Signal Processing And Control*, 68, 102577. <https://doi.org/10.1016/j.bspc.2021.102577>
- [70] Dorschky, E., Camomilla, V., Davis, J., Federolf, P., Reenalda, J., & Koelewijn, A. D. (2022b). Perspective on “in the wild” movement analysis using machine learning. *Human Movement Science*, 87, 103042. <https://doi.org/10.1016/j.humov.2022.103042>
- [71] Jiménez-Grande, D., Atashzar, S. F., Devecchi, V., Martinez-Valdes, E., & Falla, D. (2022). A machine learning approach for the identification of kinematic biomarkers of chronic neck pain during single- and dual-task gait. *Gait & Posture*, 96, 81-86. <https://doi.org/10.1016/j.gaitpost.2022.05.015>
- [72] Rahayu, E. S., Yuniarno, E. M., Purnama, I. K. E., & Purnomo, M. H. (2023). Human activity classification using deep learning based on 3D motion feature. *Machine Learning With Applications*, 12, 100461. <https://doi.org/10.1016/j.mlwa.2023.100461>
- [73] Chambers, C. C., Lynch, T. S., Gibbs, D. B., Ghodasra, J. H., Sahota, S., Franke, K., Mack, C. D., & Nuber, G. W. (2016). Superior Labrum Anterior-Posterior Tears in the

VI. REFERENCIAS

National Football League. *The American Journal Of Sports Medicine*, 45(1), 167-172. <https://doi.org/10.1177/0363546516673350>