

# UCUENCA

## Universidad de Cuenca

Facultad de Ingeniería

Carrera de Electrónica y Telecomunicaciones

**Análisis comparativo entre sistemas de clasificación basados en técnicas de Machine Learning para catalogar tipos de empuje de un patinador de velocidad.**

Trabajo de titulación previo a la  
obtención del título de Ingeniero  
en Electrónica y  
Telecomunicaciones


**Autores:**

Ximena Lorena Albornoz Tépan

Bryan Sebastián Ulloa Montaleza

**Director:**

Darwin Fabián Astudillo Salinas

ORCID:  0000-0001-7644-0270

**Cuenca, Ecuador**

2023-08-23

## Resumen

El patinaje de velocidad es un deporte de prestigio que requiere habilidades técnicas y un estado físico óptimo. Sin embargo, es necesario contar con herramientas que complementen el entrenamiento de los deportistas en formación. En este sentido, se propone el uso de sistemas de visión artificial y Machine Learning (ML) específicamente, el sistema *OpenPose*, para obtener datos sobre las articulaciones y los puntos clave del patinador. Además, se utilizan sistemas de ML para identificar y clasificar el tipo de empuje que utiliza un patinador cuando realiza técnica de recta. La investigación explora dos enfoques principales. El primero consiste en la clasificación de imágenes mediante una Convolutional Neural Network (CNN) utilizando la arquitectura Visual Geometry Group (VGG19), logrando una exactitud del 90.72%. El segundo enfoque utiliza vectores de características a través de un análisis biomecánico con un sistema Support Vector Machine (SVM) y un algoritmo Random Forest (RF), obteniendo una exactitud del 94% y 92% respectivamente. Se presentan diversas métricas de evaluación, como exactitud, precisión, sensibilidad y F1-score, junto con matrices de confusión y curvas Receiver Operating Characteristic (ROC). También se analizan características de los enfoques, como el uso del aumento de datos para combatir el sobreajuste. Se incluyen tablas con los resultados de variables biomecánicas y se muestran imágenes de los ángulos característicos durante la ejecución de las técnicas de empuje. En general, los enfoques de CNN y SVM demuestran resultados prometedores en la clasificación de datos para los diferentes tipos de empuje.

*Palabras clave:* Clasificación, CNN, SVM, *OpenPose*, Patinaje de velocidad



**El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Cuenca ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por la propiedad intelectual y los derechos de autor.**

Repositorio Institucional: <https://dspace.ucuenca.edu.ec>

### Abstract

Speed skating is a highly regarded sport that requires technical skills and an optimal physical condition; nevertheless, there is a need for a tool that complements the training of athletes in training. Thus, the use of artificial vision and Machine Learning (ML) systems is proposed. For the former, the OpenPose system is used to obtain data about the skater's joints and keypoints; the latter, to identify and classify the type of push when skaters use the inline technique. In the research, 2 main focuses are explored. The first one consists of image classification through a Convolutional Neural Network (CNN), using the Visual Geometry Group (VGG)19 architecture, achieving an accuracy of 90.72 %. The second approach uses feature vectors through a biomechanical analysis with a Support Vector Machine (SVM) system and a Random Forest (RF) algorithm, obtaining an accuracy of 94 % y 92 %, respectively. Different evaluation metrics are presented, such as accuracy, precision, sensitivity, and F1-score, along with a confusion matrix and Receiver Operating Characteristic (ROC) curves. Moreover, features of the prior-mentioned approaches are discussed, for instance, the use of data augmentation to combat overfitting. Furthermore, the study includes result tables of biomechanical variables and pictures of distinctive angles during the execution of forward pushing techniques. Overall, the CNN and SVM exhibit promising results in the classification of data for the different types of pushes in speed skating inline technique.

*Keywords:* Classification, CNN, SVM, *OpenPose*, speed skating



**El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Cuenca ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por la propiedad intelectual y los derechos de autor.**

Repositorio Institucional: <https://dspace.ucuenca.edu.ec>

## Índice general

1	Introducción	17
1.1	Identificación del problema	17
1.2	Justificación	18
1.3	Alcance	19
1.4	Objetivos	21
1.4.1	Objetivo general	21
1.4.2	Objetivos específicos	21
2	Marco teórico	22
2.1	Principios básicos del patinaje de velocidad	22
2.1.1	Ciclo del patinaje de velocidad	22
2.1.2	Técnica y fases	23
2.1.3	Tipos de empuje	25
2.1.3.1	Empuje clásico	25
2.1.3.2	Empuje pendular	26
2.1.3.3	Doble empuje	26
2.2	Modelos biomecánicos	27
2.2.1	Parámetros	28
2.2.2	Evaluación del patinaje de velocidad	29
2.2.2.1	Herramientas para la captura de movimiento	29
2.3	OpenPose	30
2.4	Tratamiento de datos	34
2.4.1	Preparación de datos	35
2.4.1.1	Normalización	35
2.4.1.2	Interpolación	36
2.4.1.3	Técnicas de suavizado de datos	36

2.4.2	Imagen de energía de la marcha (GEI) . . . . .	38
2.4.2.1	Distancia entre dos puntos en dos dimensiones . . . . .	40
2.4.2.2	Ángulo formado por tres puntos . . . . .	40
2.5	Algoritmos de clasificación . . . . .	40
2.5.1	Redes neuronales convolucionales (CNN) . . . . .	41
2.5.1.1	Modelo <i>Visual Geometry Group</i> (VGG19) . . . . .	42
2.5.2	Máquinas de vectores de soporte (SVM) . . . . .	44
2.6	Evaluación de algoritmos de clasificación . . . . .	47
2.6.1	Matriz de confusión . . . . .	47
2.6.2	Curva ROC . . . . .	48
2.6.3	Otras métricas de evaluación . . . . .	49
2.6.3.1	Exactitud . . . . .	49
2.6.3.2	Precisión . . . . .	50
2.6.3.3	Sensibilidad . . . . .	50
2.6.3.4	F1-score . . . . .	50
3	Trabajos relacionados . . . . .	51
3.1	Human Action Recognition (HAR) y aprendizaje automático en el ámbito deportivo . . . . .	52
3.2	OpenPose como herramienta de HAR . . . . .	53
3.3	Gait Energy Image (GEI) como un enfoque de Computer Vision (CV) en diferentes aplicaciones . . . . .	55
3.4	Construcción de conjuntos de datos . . . . .	56
3.5	Aplicación de técnicas de aprendizaje automático para la clasificación de técnicas y movimientos en deportes . . . . .	58
4	Metodología . . . . .	61
4.1	Protocolo de adquisición de datos . . . . .	61
4.1.1	Distribución de equipos . . . . .	62
4.1.2	Descripción de Hardware . . . . .	63
4.1.2.1	Descripción de los equipos de captura de video . . . . .	64
4.1.3	Herramientas informáticas . . . . .	64
4.2	Pre-procesamiento de datos . . . . .	65

4.2.1	Corte y sincronización de videos . . . . .	66
4.2.2	Extracción del esqueleto . . . . .	66
4.2.3	Corrección de errores . . . . .	68
4.2.4	Suavizado de datos . . . . .	75
4.2.5	Normalización de datos . . . . .	76
4.3	Extracción de las características biomecánicas . . . . .	79
4.3.1	Obtención del Skeleton Gait Energy Image (SGEI) . . . . .	79
4.3.2	Cálculo de variables espacio-temporales . . . . .	81
4.3.3	Cálculo de ángulos . . . . .	83
4.4	Métodos de clasificación . . . . .	85
4.4.1	Clasificación con CNN . . . . .	86
4.4.2	Clasificación con SVM . . . . .	87
4.4.3	Clasificación con RF . . . . .	89
5	Resultados . . . . .	90
5.1	Análisis . . . . .	90
5.1.1	Parámetros espacio-temporales y ángulos . . . . .	90
5.1.2	Comparación en los ángulos de las extremidades inferiores . . . . .	94
5.1.3	Evaluación de los clasificadores . . . . .	97
5.1.3.1	Clasificación usando CNN . . . . .	98
5.1.3.2	Clasificación usando SVM . . . . .	100
5.1.3.3	Clasificación usando RF . . . . .	102
5.1.3.4	Comparación entre los métodos de clasificación . . . . .	103
6	Conclusiones y recomendaciones . . . . .	106
6.1	Conclusiones . . . . .	106
6.2	Recomendaciones . . . . .	108
6.3	Trabajos futuros . . . . .	108
	Referencias . . . . .	110
	Apendices . . . . .	120
A	<i>Openpose</i> . . . . .	121

A.1	Instalación . . . . .	121
A.2	Ejecución . . . . .	122
B	Valores incluidos en los vectores de características	123
B.1	Máximos, mínimos y valor medio de los ángulos (promedio $\pm$ desviación estándar) . . . . .	123
Anexos		125
A	Permiso	125
A.1	Permiso Club deportivo . . . . .	125
B	Consentimiento informado	126
B.1	Consentimiento informado menor de edad . . . . .	126
B.2	Consentimiento informado mayor de edad . . . . .	127
C	Registro de firmas	128
C.1	Menores de edad . . . . .	128
C.2	Mayores de edad . . . . .	129

## Índice de figuras

2.1 Fases de la recta (ciclo derecho) . . . . .	24
2.2 Empuje clásico (ciclo derecho) . . . . .	25
2.3 Empuje pendular (ciclo derecho) . . . . .	26
2.4 Doble empuje (ciclo derecho) . . . . .	27
2.5 Extracción de la pose en <i>OpenPose</i> . . . . .	30
2.6 Esquema de funcionamiento de <i>OpenPose</i> . Fuente: [1] . . . . .	31
2.7 Arquitectura de <i>OpenPose</i> . Fuente: [1, 2] . . . . .	32
2.8 Formato de salida de pose (BODY_25). Fuente: [3] . . . . .	34
2.9 Ejemplo de una ventana de <i>Hanning</i> con N=50 . . . . .	37
2.10 Curva antes y después del suavizado. . . . .	38
2.11 <i>Gait Energy Image</i> GEI. Fuente: [4]. . . . .	39
2.12 <i>Skeleton Gait Energy Image</i> (SGEI). . . . .	39
2.13 Modelo 19. Fuente: [5] . . . . .	44
2.14 Clasificación con SVM. Fuente: [6]. . . . .	45
2.15 Uso del <i>kernel trick</i> para introducir datos en un espacio de característi- cas de alta dimensionalidad. Fuente: [7]. . . . .	46
2.16 Curva ROC. Fuente:[8]. . . . .	48
2.17 Área bajo la curva ROC. Fuente: [8]. . . . .	49
4.1 Distribución de las cámaras y marcas de referencia. . . . .	62
4.2 Esquema de procesos. . . . .	65
4.3 Fotograma de un video analizado con el esqueleto sobrepuesto sobre la persona. . . . .	67
4.4 Contenido del archivo JSON para el fotograma analizado. . . . .	67
4.5 Diagrama de flujo para la obtención de la serie temporal. . . . .	69
4.6 Series temporales para las coordenadas en $x$ de los puntos de los tobi- llos, Vista Sagital. . . . .	70



4.7	Serie temporal para la coordenada en $x$ del tobillo derecho, Vista Sagital.	71
4.8	Ubicación de los saltos correspondientes a errores en la serie temporal.	72
4.9	Separación en intervalos sin problemas e intervalos con errores. . . . .	73
4.10	Diagrama de flujo para corrección de errores mediante suavizado iterativo.	73
4.11	Suavizado iterativo para la corrección de errores. . . . .	74
4.12	Suavizado iterativo para la corrección de errores en la serie temporal. .	75
4.13	Serie temporal con y sin suavizado. . . . .	76
4.14	Normalización de datos en la vista frontal. . . . .	77
4.15	Serie temporal correspondiente a la coordenada en $x$ del tobillo derecho, Vista frontal. . . . .	78
4.16	Secuencia de siluetas binarias frontales. . . . .	80
4.17	Secuencia de siluetas binarias sagitales. . . . .	80
4.18	Imágenes SGEI generadas a partir de siluetas binarias . . . . .	81
4.19	Indicación de las ubicaciones en las que se miden las variables espaciotemporales. . . . .	83
4.20	Ángulos calculados en vista frontal. . . . .	85
4.21	Ángulos calculados en vista sagital. . . . .	85
4.22	Diagrama de flujo para la clasificación con CNN. . . . .	86
4.23	Diagrama de flujo para la clasificación con SVM y RF. . . . .	88
5.1	Flexión del tronco . . . . .	95
5.2	Flexión de rodilla izquierda. . . . .	96
5.3	Flexión de tobillo derecho. . . . .	97
5.4	Curvas de entrenamiento. . . . .	99
5.5	Evaluación de CNN. . . . .	100
5.6	Evaluación de SVM. . . . .	102
5.7	Métodos de evaluación. . . . .	103

### Índice de tablas

2.1	Formatos de salida de <i>OpenPose</i> . . . . .	33
2.2	Orden de mapeo de partes del cuerpo para el modelo <i>BODY25</i> . . . . .	33
2.3	Modelo 19. Fuente: [9] . . . . .	43
2.4	Matriz de confusión para una clasificación binaria . . . . .	48
4.1	Ángulos para la vista frontal . . . . .	84
4.2	Ángulos para la vista sagital . . . . .	84
5.1	Distancias, tiempos y velocidades (promedio $\pm$ desviación estándar) . .	92
5.2	Valores medios de los ángulos en vista frontal (promedio $\pm$ desviación estándar) . . . . .	93
5.3	Valores medios de los ángulos en vista sagital (promedio $\pm$ desviación estándar) . . . . .	93
5.4	Evaluación de diferentes métricas para CNN. . . . .	99
5.5	Evaluación de diferentes métricas para SVM. . . . .	101
5.6	Evaluación de diferentes métricas para RF . . . . .	102
5.7	Características de los sistemas de clasificación. . . . .	105
B.1	Ángulos en vista frontal completos . . . . .	123
B.2	Ángulos en vista sagital completos (promedio $\pm$ desviación estándar) . .	124

**Dedicatoria**

Dedico este trabajo a toda mi familia, por el apoyo y confianza depositados en mí, por estar en cada paso. Sobre todo quiero dedicar este trabajo a mis padres Orlando y Mercedes por su paciencia. A mis sobrinos José y Sebastián que son mi razón de seguir adelante, cada página de este trabajo es de ustedes también, por ser mi inspiración y mi motor en momentos difíciles, espero que algún día puedan entender lo importantes que son para mí.

**Ximena Albornoz**

**Dedicatoria**

Dedico este trabajo de tesis a mi querida familia, cuyo amor y apoyo incondicional han sido la fuerza impulsora detrás de mi éxito académico. Agradezco profundamente la ayuda y cariño que me han brindado mi madre Guadalupe y mi padre Vicente.

**Bryan**

### Agradecimientos

Me gustaría agradecer a Dios, a mis padres, amigos y a todas las personas maravillosas que me han apoyado a lo largo de estos años. Toda su fe en mí y apoyo incondicional han sido fundamentales para mi crecimiento y desarrollo.

A mis padres, gracias por siempre estar a mi lado, por su amor incondicional. Por su guía y darme las herramientas necesarias para afrontar la vida. Nunca podre agradecerles lo suficiente.

A Gabriela Cavaliere, Micaela Matute y a todo el club de patinaje Gabriela Cavaliere, por permitirme ser parte del club, por su tiempo y apoyo en todo momento en el desarrollo de este trabajo. Toda su generosidad y amabilidad ha dejado una huella en mí.

A mis amigos, Sebastián Zhindon, Nathaly Guailas, Angélica Ordóñez, Bryan Martínez, Delsa Silva, Markos Chica, Fernando Malla, Fernando Remache, Emilia Chica, Joel Guaman, Paul Tigre, Jarni flores y David Cárdenas. Gracias por todos los momentos inolvidables, las risas, las aventuras y por haber hecho mi paso por la universidad un lugar más agradable. Por estar en los momentos difíciles, muchas gracias. Han hecho mi vida mejor de muchas formas.

A mi compañero de tesis y colega, Sebastián Ulloa, por la dedicación, empeño, comprensión y paciencia puesto en este trabajo. De igual manera quisiera agradecer a mi tutor Ing. Fabián Astudillo, por brindarnos los conocimientos necesarios y por habernos impulsado a realizar este trabajo. Gracias por su interés genuino en nuestro éxito y por su disposición en ayudarnos en cualquier momento.

Gracias a cada uno de ustedes por creer en mí, por ser parte de mi vida, por el impulso a ser cada día mejor. Todo su apoyo ha hecho posible que alcance mis sueños. Sin ustedes, no sería quien soy. Y sobre todo quiero agradecer a Dios, por estar en cada paso de mi vida, por la fortaleza, sabiduría, amor y guía que me brindaste.

Gracias, de todo corazón.

**Ximena Albornoz**

### Agradecimientos

Quiero agradecer a mi familia por su apoyo y cariño incondicional durante todo mi proceso de estudios, a mi madre quien siempre vela por mí, a mi padre que me ha apoyado siempre y a mis hermanos que me han ayudado cuando lo necesitaba, ellos me han ayudado a ser quien soy hoy en día.

A mis compañeros y amigos Nathaly Guallas, Paul Tigre, Joel Guaman, Jae Hwang, Jarni Flores por haber compartido un valioso tiempo tanto dentro como fuera del salón de clases y por todas las experiencias que hemos pasado juntos. A Sebastián Zhindon por habernos brindado gentilmente su ayuda.

A mis amigos Jaime Iñiguez y Christopher Ramón por haberme apoyado en los momentos más difíciles, por haberme escuchado y aconsejado.

A mi compañera de tesis Ximena Albornoz, tanto por su amistad como por haber sido siempre alguien en quien confiar y con quien trabajar como colegas.

Al Ing. Fabián Astudillo por brindarnos su ayuda y conocimientos en el desarrollo de este trabajo, así como por su labor de docente.

Les agradezco de corazón por todas las experiencias vividas y por todo lo que he aprendido gracias a ellos.

**Bryan**

## Abreviaciones y acrónimos

**ADAM** Adaptive Moment Estimation. 56

**AT** Attention Mechanisms. 58

**AUC** Area Under the ROC Curve. 49, 52, 101, 103

**CEDIA** Corporación Ecuatoriana para el Desarrollo de la Investigación y la Academia.  
63, 64

**CNN** Convolutional Neural Network. 2, 3, 6, 9, 10, 19, 30, 31, 38, 41, 52–56, 58–60,  
66, 79, 86, 87, 97–100, 103–107

**CV** Computer Vision. 5, 51, 53, 55, 106

**DL** Deep Learning. 34, 51, 52, 58

**FCNN** Fully Connected Neural Network. 58, 59

**FPR** False Positive Rate. 48

**GEI** Gait Energy Image. 5, 8, 38, 39, 55–57, 60

**HAR** Human Action Recognition. 5, 51–53, 56

**HD** High Definition. 64

**IMU** Inertial Measurement Unit. 58, 59

**KNN** K Nearest Neighbors. 52, 53, 55, 58, 59

**LSTM** Long Short-Term Memory. 52–54, 56, 59, 60

**MEI** Motion Energy Image. 56, 57

**MHI** Motion History Image. 56, 57

**MIG** Multi-Instance GPU. 63

**ML** Machine Learning. 2, 34, 38, 41, 44, 51–53, 57, 58, 60, 61, 87, 106, 107

**MOCAP** Motion Capture. 18, 29

**PAF** Part Affinity Fields. 30–32

**PCA** Principal Component Analysis. 53

**RAM** Random Access Memory. 63, 108

**RBF** Radial Basis Function. 59, 89, 100, 101

**ResNet** Residual Network. 56

**RF** Random Forest. 2, 6, 9, 10, 53, 54, 59, 87–89, 102–107

**RGB** Red Green Blue. 30, 43, 58

**RNN** Recurrent Neural Network. 54

**ROC** Receiver Operating Characteristic. 2, 5, 8, 47–49, 97, 99, 101, 103, 105, 107

**SGD** Stochastic Gradient Descent. 41, 98

**SGEI** Skeleton Gait Energy Image. 6, 9, 39, 65, 78–81, 86, 98, 99, 106, 107

**SVM** Support Vector Machine. 2, 3, 6, 8–10, 19, 44–46, 52, 53, 58–60, 79, 87–89, 97, 100–104, 106, 107

**TPR** True Positive Rate. 48

**UHD** Ultra High Definition. 64

**VGG19** Visual Geometry Group. 2, 42, 56, 86, 107

**VRAM** Video Random Access Memory. 63, 108



## Capítulo 1 Introducción

En este capítulo, se abordan distintos aspectos relacionados con el presente trabajo de titulación. En la Sección 1.1 la identificación del problema. Posteriormente, en la Sección 1.2 se expone la justificación que respalda la relevancia de este trabajo. A continuación, en la Sección 1.3 da a conocer el alcance del trabajo. Finalmente, la Sección 1.4 presenta el objetivo general y los objetivos específicos planteados. Cada una de estas secciones contribuye a proporcionar una comprensión completa del contexto y los propósitos de este estudio.

### 1.1. Identificación del problema

En la actualidad, el patinaje es uno de los deportes más prestigiosos en el ámbito competitivo, tanto en eventos nacionales como internacionales. Para sobresalir en esta disciplina, los patinadores requieren de una combinación de habilidades técnicas y un óptimo estado físico, lo que les permite alcanzar su máximo rendimiento sobre la pista [10]. Con el fin de medir objetivamente el desempeño de los patinadores, es fundamental comprender la biomecánica involucrada en su práctica [11]. A pesar de los avances tecnológicos en el ámbito deportivo en los últimos años, aún no se dispone de una herramienta de uso sencillo que permita a los entrenadores obtener una retroalimentación precisa sobre la técnica de cada patinador.

Para los monitores de disciplinas deportivas, es fundamental realizar un seguimiento y control con el objetivo de mejorar su técnica y prevenir lesiones. Actualmente, existen diferentes opciones disponibles para abordar estas problemáticas, como el uso de tecnologías multimedia empleadas en el análisis de imágenes y estudio de videos deportivos con el fin de encontrar fallas y perfeccionar la técnica [12]. Una de estas herramientas es el software Kinovea, ampliamente utilizado por entrenadores para analizar el rendimiento de los deportistas [13]. Sin embargo, este programa es genérico y no está adaptado específicamente a cada deporte, por lo que implica que los puntos claves del cuerpo y ángulos deseados deben ser seleccionados de forma manual, lo cual es un proceso que consume mucho tiempo y resulta poco eficiente.

En [14] se ha desarrollado un sistema para evaluar el movimiento de los bateado-

res en el deporte de Béisbol, con el objetivo de corregir sus posturas y mejorar su rendimiento. Esta iniciativa tecnológica ha impulsado el estudio de la corrección postural en diferentes disciplinas deportivas, buscando alcanzar un rendimiento óptimo en los atletas. Por ejemplo, en el estudio [15] se demuestra la viabilidad de evaluar el rendimiento en el patinaje de velocidad. En su investigación utilizan la tecnología de Motion Capture (MOCAP) junto con un software avanzado que permite interpretar el movimiento humano a modelos biomecánicos, logrando un análisis detallado y preciso del rendimiento en esta disciplina.

La retroalimentación proporcionada por estos sistemas ofrece a los patinadores una comprensión más profunda de su técnica y les permite mejorar su rendimiento. Los entrenadores tienen la opción de utilizar esta tecnología para identificar y corregir problemas específicos en la técnica de sus atletas. En el contexto del patinaje de velocidad, el uso de tecnología de MOCAP tiene un gran potencial para mejorar tanto el rendimiento como la técnica de los patinadores.

Debido a los puntos expuestos anteriormente, se observa la necesidad de contar con una herramienta tecnológica que facilite las correcciones y el método de entrenamiento por parte de los monitores. Esta herramienta permitiría optimizar el proceso de formación y brindaría un enfoque más efectivo para el entrenamiento de los patinadores.

## 1.2. Justificación

La evaluación de la técnica en el patinaje de velocidad es crucial para mejorar el rendimiento de los deportistas en competencias [15]. Actualmente, una técnica comúnmente empleada para evaluar el patinaje implica la grabación de secuencias de video seguidas de una revisión manual por parte de expertos para analizar los movimientos del patinador. Sin embargo, este proceso demanda considerable consumo de tiempo [16].

En el marco de la investigación científica actual, se ha observado un creciente número de artículos y estudios que exploran los beneficios de emplear técnicas de visión artificial y *Machine Learning* en el ámbito deportivo. Entre los enfoques de la visión artificial utilizados en esta área, destaca el reconocimiento del cuerpo humano y sus

partes clave, como articulaciones y extremidades [14, 17]. Estos enfoques se basan en el análisis de material audiovisual como punto de partida para extraer las características del movimiento y la postura de los deportistas.

Por otro lado, se tiene *Machine Learning*, que se ha convertido en una herramienta poderosa para la clasificación de patrones en diversos campos. Estos sistemas tienen la capacidad de analizar grandes conjuntos de datos, extraer patrones y características que resultan difíciles de detectar para los seres humanos. En el ámbito del deporte, se emplean técnicas de *Machine Learning*, como las SVM, para procesar y clasificar datos obtenidos a través de sistemas de visión artificial [18], así como para generar modelos predictivos del desempeño de los atletas [19]. Además, las CNN son utilizadas para extraer características a partir de imágenes [20, 21].

En este contexto, el presente estudio propone el uso del sistema de visión artificial *OpenPose* para obtener las coordenadas bidimensionales de las articulaciones y puntos clave de un patinador mientras se desplaza por la pista. Estas coordenadas se utilizan como datos de entrada en sistemas *Machine Learning* los cuales se encargan de identificar y clasificar el tipo de empuje que el patinador utiliza durante el desplazamiento. El objetivo principal es evaluar el rendimiento de diversos sistemas de *Machine Learning* en la identificación y clasificación de los tipos de empuje de un patinador de velocidad a partir de secuencias de video. Esto permitirá al entrenador obtener una mejor percepción de cada deportista y su técnica.

Estos resultados presentan un gran potencial de aplicación en diferentes áreas prácticas, como el deporte y la medicina, lo cual destaca la importancia de investigar la biomecánica en el patinaje y emplear tecnologías y herramientas innovadoras para su perfeccionamiento.

### 1.3. Alcance

Este proyecto implica el desarrollo de una solución integrada que combina sistemas de visión artificial y *Machine Learning* para analizar la técnica en el patinaje de velocidad. Este sistema no invasivo, permite medir una variedad de parámetros biomecánicos y tiene como objetivo diferenciar las técnicas de patinaje utilizadas por un patinador.

Para lograr esto, se ha llevado a cabo una investigación exhaustiva sobre los modelos biomecánicos específicos del patinaje de velocidad, en conjunto con la aplicación de técnicas avanzadas de *Machine Learning* para clasificar las distintas técnicas de patinaje.

En una primera etapa, se extraen las coordenadas  $(x, y)$  de los puntos clave del cuerpo humano a partir de archivos de video. Posteriormente, se aplican funciones de interpolación y suavizado con el fin de minimizar los errores de estimación de coordenadas. Para lograr un análisis preciso, se realiza una caracterización del esqueleto, es decir, que sea independiente de la posición en la que se encuentre y se obtienen los parámetros espacio-temporales necesarios. Estos parámetros incluyen la longitud del paso, el tiempo de paso, la longitud de zancada, la velocidad de empuje, la dirección de empuje, la inclinación de la postura frontal, así como los ángulos formados por las extremidades del cuerpo humano.

Con los datos obtenidos, se comparan diferentes técnicas de *Machine Learning* para clasificar el tipo de técnica (se consideran 3 tipos) que el patinador está empleando. Este enfoque permite detectar patrones y tendencias en el movimiento de los patinadores, además de desarrollar métodos para mejorar su técnica y rendimiento en la pista de patinaje. Una vez obtenidos los parámetros espacio-temporales y realizada la clasificación de la técnica del patinador, se procede a evaluar el rendimiento y la eficacia del sistema de clasificación implementado. Para ello, se llevan a cabo pruebas y experimentos destinados a determinar la precisión y confiabilidad del sistema en la detección y clasificación de la técnica de patinaje.

En resumen, la implementación de un sistema de visión e inteligencia artificial en el análisis de la biomecánica del patinaje permite obtener parámetros espacio-temporales precisos y evaluar la clasificación de la técnica utilizada por los patinadores. Estas aplicaciones tienen un potencial significativo tanto en el ámbito deportivo, como herramienta de entrenamiento y mejora de la técnica de patinaje, así como en el ámbito médico, como herramienta de diagnóstico, seguimiento de lesiones y recuperación de patinadores lesionados.

## 1.4. Objetivos

### 1.4.1. Objetivo general

Evaluar sistemas de clasificación basados en *Machine Learning* para catalogar los tipos de empuje de un patinador de velocidad a partir de secuencias de video.

### 1.4.2. Objetivos específicos

El presente trabajo de titulación tiene los siguientes objetivos específicos:

- Estudiar los modelos biomecánicos del empuje en el patinaje de velocidad y el estado del arte del uso de *Machine Learning* en el análisis de actividades deportivas.
- Crear un conjunto de datos de entrenamiento, validación y evaluación para los sistemas de clasificación a partir de secuencias de video para los 3 tipos de empuje en el patinaje de velocidad.
- Implementar al menos 2 clasificadores basados en técnicas de *Machine Learning* y comparar los resultados de clasificación del tipo de empuje.

## Capítulo 2 Marco teórico

El presente capítulo está estructurado en seis secciones, las cuales se describen a continuación. En la Sección 2.1 se presentan los principios básicos del patinaje de velocidad, incluyendo conceptos fundamentales y técnicas utilizadas en esta disciplina deportiva. La Sección 2.2 se enfoca en los modelos biomecánicos aplicados al patinaje de velocidad, describiendo cómo se estudia la mecánica del movimiento en esta actividad. En la Sección 2.3 se introduce *OpenPose*, una herramienta de estimación de pose ampliamente utilizada en la visión artificial, que permite analizar y capturar el movimiento de los patinadores en dos dimensiones. En la Sección 2.4 se aborda el tratamiento de datos, incluyendo técnicas de procesamiento utilizadas para mejorar la precisión del análisis en el patinaje de velocidad. Los algoritmos de clasificación para identificar patrones de movimiento específicos se presentan en la Sección 2.5, mientras que en la Sección 2.6 se describen las herramientas utilizadas para evaluar la efectividad de dichos algoritmos.

### 2.1. Principios básicos del patinaje de velocidad

El patinaje de velocidad es un deporte que involucra una combinación de aspectos físicos, técnicos y estratégicos. Se basa en el deslizamiento lateral y diagonal sobre los patines, una habilidad que se desarrolla a partir del patrón de caminar y correr [22]. Además, se compone de una serie de posturas que producen un equilibrio y posición corporal óptima para maximizar la eficiencia en el desplazamiento [23].

Este deporte es de tipo cíclico, debido a que el propio cuerpo es el motor y transmisor de la actividad motora. El patín, como medio mecánico, implica la ejecución de movimientos que pueden resultar antinaturales. Por lo tanto, es importante desarrollar y aplicar una técnica eficiente que permita optimizar la capacidad condicional y convertirla en un movimiento eficaz y fluido [23, 24].

#### 2.1.1. Ciclo del patinaje de velocidad

El patinaje es una habilidad en constante movimiento que se compone de ciclos periódicos de patrones de movimiento [25]. Estos ciclos pueden cambiar por distintos

parámetros condicionales y coordinativos entre los movimientos de las extremidades inferiores y superiores. Además, se requiere una distribución equilibrada del peso corporal para lograr la estabilidad en el patín y la generación de la fuerza necesaria para impulsarse [26]. Estos factores contribuyen a definir un estilo de ejecución particular. Es así que, el análisis del movimiento involucra conocer las distintas fases o componentes que lo conforman, así como las características cualitativas y cuantitativas del mismo. No obstante, la invariabilidad y constancia en la ejecución de los ciclos de movimiento depende en gran medida de la resistencia física y la destreza técnica del patinador [24, 27].

### **2.1.2. Técnica y fases**

Por técnica se entiende una adecuada y económica forma de trasladarse sobre los patines, manteniendo una postura ideal y realizando movimientos coordinados, con el fin de obtener los mejores resultados deportivos [28, 29]. Consiste en una serie de movimientos simétricos y continuos, en cuanto al desplazamiento en recta, requiriendo una posición básica sólida [24]. Cada patinador tiene características antropométricas propias, las mismas que deben ser consideradas en la técnica del patinaje. La posición básica estática de un deportista en esta área consiste en una combinación de movimientos, como mantener una posición baja, con una flexión de rodillas entre  $90^\circ$  y  $110^\circ$ , un adecuado giro de tobillo, movimiento de cadera y un trabajo adecuado de las extremidades superiores para mantener el equilibrio [24, 26]. Asimismo, se enfoca en lograr una distribución correcta del peso corporal, lo que permite una ejecución fluida y una transición suave entre las diferentes fases del ciclo de patinaje [30].

El patinaje cuenta con cuatro tipos de técnicas: técnica de salida, técnica de curva, técnica de recta y técnica de llegada. Cada una de ellas consta de varias fases que deben seguirse. En el caso de la técnica de recta, se debe cumplir los siguientes pasos (consultar Figura 2.1).



Figura 2.1: Fases de la recta (ciclo derecho)

**Fase 1. Posición Básica:** en esta fase, el patín que se está recuperando se ubica al lado del patín opuesto, es decir, tiene un doble apoyo. La separación entre los patines debe ser menor que la distancia entre los hombros. Además, es importante centrar el peso del cuerpo sobre el borde del patín deslizante. El brazo contrario al patín que realizará el empuje debe estar hacia afuera, mientras que el brazo lateral hacia adentro [15, 31].

**Fase 2. Empuje:** es cuando el deportista ejerce una fuerza lateral y hacia la parte posterior para generar un desplazamiento, tomando en cuenta que es necesario tener la máxima amplitud con el contacto continuo del patín [30, 32]. La extremidad de la rodilla del patín de apoyo debe formar un ángulo cercano a  $90^\circ$ . Durante esta fase el peso del cuerpo se distribuye entre el pie y el patín de apoyo. Además, el brazo opuesto al empuje se mueve hacia arriba y atrás, mientras que el brazo del mismo lado se mueve hacia adelante [15]. En cuanto a esta fase, existen tres tipos de empuje que el deportista puede realizar, el empuje clásico, el empuje pendular y el doble empuje.

**Fase 3. Despegue:** esta fase empieza con un impulso final. La extremidad inferior está en su máxima extensión, por lo tanto, el patín realiza un despegue del suelo desde la parte posterior [32]. Durante esta fase, el peso corporal se encuentra sobre la rodilla y la parte externa del patín de apoyo [15].

**Fase 4. Recuperación:** Durante esta fase, el patín que se despegó es llevado hacia la parte posterior del cuerpo en un movimiento de recobro hacia adentro [15];



aquí se inicia el primer desplazamiento. Es importante mantener un ángulo óptimo en la rodilla de la pierna de apoyo [25, 30].

**Fase 5. Recobro:** se lleva a cabo un recobro hacia adelante realizando un desplazamiento que retorna a la posición básica, mientras el patinador se prepara para realizar el empuje con la pierna contraria. Al utilizar el patín para recuperarse, se inicia un nuevo movimiento de impulso al adoptar una postura vertical en la parte lateral del patín. Esto marca el inicio de un ciclo de movimiento completamente distinto [15].

### 2.1.3. Tipos de empuje

En el patinaje de velocidad, el empuje hace referencia a la fuerza que un patinador ejerce contra la superficie de la pista para impulsarse en la dirección de su trayectoria. La dirección del empuje puede variar según la velocidad del patinador y la inclinación de la superficie [32].

#### 2.1.3.1. Empuje clásico

En el empuje clásico, se ejerce una fuerza de impulso lateral. El patín de apoyo se mantiene perpendicular al suelo y requiere mantener una alineación con el tobillo, rodilla, cadera y hombro [33, 34]. Esta técnica es muy precisa y solo permite ajustar la flexión de las rodillas y la frecuencia de zancada, mientras el pie contrario sostiene el peso del cuerpo [32] (ver Figura 2.2).

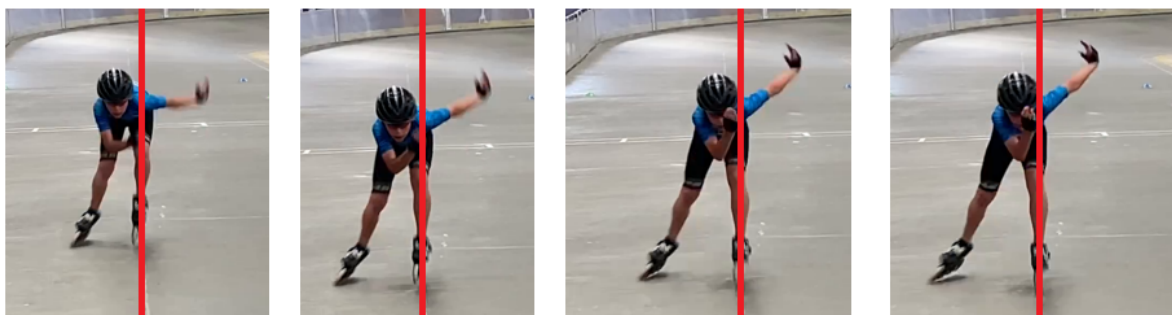


Figura 2.2: Empuje clásico (ciclo derecho)

### 2.1.3.2. Empuje pendular

El empuje pendular es cuando un patín realiza el empuje, mientras que el patín contrario ejerce una tracción con una inclinación de aproximadamente  $\pm 25^\circ$ , lo que resulta en una mayor aceleración. Esto, a su vez, desplaza el centro de gravedad hacia el exterior del patín. Sin embargo, el empuje y la tracción no se llevan a cabo simultáneamente [25, 29, 33] (ver Figura 2.3).



Figura 2.3: Empuje pendular (ciclo derecho)

### 2.1.3.3. Doble empuje

Durante el doble empuje, se produce un movimiento constante en la pierna de apoyo, donde se realiza una transición suave entre el empuje y la tracción. Estas acciones se llevan a cabo simultáneamente, lo que resulta en un menor gasto de energía y un aumento en la fuerza de reacción [32, 33]. A medida que se produce esta fase de empuje, el cuerpo se desplaza lateralmente con una inclinación de aproximadamente  $\pm 25$  grados [25]. El empuje se realiza hacia adentro previo al empuje lateral; además, la posición del cuerpo varía, debido a que, el cuerpo se desplaza en la misma dirección que el empuje interno durante la recuperación, aprovechando la transferencia del peso corporal hacia el pie contrario [33] (ver Figura 2.4).



Figura 2.4: Doble empuje (ciclo derecho)

## 2.2. Modelos biomecánicos

Los modelos biomecánicos son una herramienta útil para comprender y mejorar el rendimiento deportivo. En el caso del patinaje de velocidad, la biomecánica se enfoca en analizar el movimiento del patinador con el fin de identificar las variables que influyen en su velocidad y eficiencia. Sin embargo, es importante tener en cuenta que las características individuales del deportista dotan de singularidad y personalidad a la técnica deportiva, por lo cual no existe un modelo matemático universalmente aplicable [25]. En otras palabras, en el ámbito deportivo la técnica perfecta no existe, debido a que se establece a partir de parámetros de rendimiento mecánico que consideran factores genéticos, antropométricos y leyes físicas.

A diferencia de otras técnicas, este deporte se distingue por su técnica especial de propulsión, que permite alcanzar velocidades superiores a 50 km/h. Entre los modelos referenciales en el patinaje, se han clasificado tres técnicas diferentes: la clásica, la pendular y la doble empuje. Se diferencian por la inclinación del pie de apoyo. El empuje se realiza en un ángulo de aproximadamente  $\pm 45^\circ$  en dirección opuesta al patín que se desliza hacia adelante. Este movimiento produce una extensión breve, pero intensa de la rodilla y la cadera, tanto el plano frontal como el sagital. Además, en el patinaje de velocidad, la fuerza de reacción generada por el empuje está orientada hacia atrás, siguiendo el principio de acción y reacción [25].

El sistema de análisis biomecánico BIOMIN-PATIN se utiliza para establecer un modelo biomecánico en el patinaje; permitiendo analizar el movimiento técnico durante esta actividad [15]. Se basa en un enfoque práctico y científico que establece una

interdisciplinaridad y una transdisciplinaridad para considerar factores integrados que se dividen en tres módulos. El morfo-estructural donde se consideran variables antropométricas de cada deportista, teniendo en cuenta sus características físicas y estructurales. El funcional, que hace alusión a las propiedades de la capacidad neuromotora del deportista, las cuales son evaluadas y analizadas en detalle. Por último, el dinámico, que analiza el movimiento desde una perspectiva cualitativa y cuantitativa, abarcando aspectos cinemáticos y cinéticos.

### 2.2.1. Parámetros

Las variables espacio-temporales son elementos fundamentales que involucran medidas de espacio y tiempo en un conjunto de datos. Estos parámetros son obtenidos mediante la transformación de datos de entrada para generar una variable de salida en específico [35]. Estas variables se utilizan en sistemas de clasificación de datos. Por ejemplo, en [36] se obtienen distancias y ángulos a partir de coordenadas de entrada. De ahí que, es necesario conocer las fórmulas más utilizadas en el cálculo de las características espacio-temporales para su correcta aplicación.

Durante el análisis cinemático del movimiento de un cuerpo, se suelen considerar variables como las velocidades, los ángulos y distancias, entre otras, sin considerar las fuerzas que provocan dicho movimiento [32]. En [24, 29] destacan ciertos parámetros espacio-temporales, entre los que se incluyen la distancia entre patines, longitud y frecuencia del paso, longitud de zancada, dirección y velocidad de empuje. Los ángulos esenciales en el patinaje de velocidad se dan en tres articulaciones importantes, estos ángulos son, la articulación de la cadera, la rodilla y el tobillo. Especialmente, el ángulo del tobillo adquiere gran importancia, ya que garantiza un apoyo estable durante todo el gesto técnico [32]. Cabe destacar que los resultados de estos parámetros varían en función de la estatura, peso, flexibilidad, fuerza y coordinación del patinador [25].

### 2.2.2. Evaluación del patinaje de velocidad

Un análisis de manera visual, es una estimación cinemática del movimiento, debido a su facilidad se convierte en la aproximación más frecuente. Este tipo de análisis se basa en la preparación de todos los participantes involucrados en las pruebas. Además de la grabación y el registro de los mismos. Posteriormente, se procede a dividir los ciclos de movimiento, y finalmente realizar un análisis cualitativo y cuantitativo para definir resultados y recomendaciones [37].

Un análisis cualitativo es la observación sistemática de la calidad del movimiento, y encontrar herramientas y estrategias que contribuyan al mejoramiento del rendimiento. Este enfoque implica la evaluación de fases, descripción y eficiencia del movimiento, así como la identificación de errores y fortalezas técnicas [38]. Por otro lado, el estudio cuantitativo implica asignar valores numéricos y magnitudes a las variables cinéticas y cinemáticas, que permiten dar un análisis más preciso y detallado del comportamiento biomecánico del gesto deportivo. En este tipo de estudio, es fundamental enfocarse únicamente en las variables más importantes y significativas para el análisis, mediante el uso de herramientas avanzadas para procesar y cuantificar de manera precisa los movimientos [37, 38].

#### 2.2.2.1. Herramientas para la captura de movimiento

Hoy en día existen varias maneras de evaluar las ejecuciones técnicas de los deportistas y analizar el movimiento en términos de distancias lineares y angulares, tiempo transcurrido, visualización de movimientos simultáneos desde diferentes ángulos o planos, entre otros. Una de estas formas de evaluación es mediante el uso de sensores inerciales ubicados en la superficie corporal del deportista. El sistema Motion Capture (MOCAP) permite pasar el movimiento humano a modelos biomecánicos, tanto bidimensionales como tridimensionales mediante una combinación de *hardware* y *software* especializado [15].

El proceso MOCAP varía según el sistema utilizado y la ubicación de los marcadores en puntos anatómicos. Una vez adquiridos los datos, estos corresponden a cada punto del cuerpo y se representan en un sistema cartesiano  $(x, y, z)$ , lo que proporciona

información sobre la posición lineal de los puntos, así como otras variables derivadas [15, 37].

### 2.3. OpenPose

*OpenPose* es un sistema de visión por computadora de código abierto desarrollado por *CMU-Perceptual Computing Lab* [39]. Este sistema es capaz de detectar y rastrear los cuerpos humanos en tiempo real a través de cámaras Red Green Blue (RGB), así como procesar imágenes y videos grabados previamente. Entre las técnicas clave que emplea *OpenPose* se encuentran las Convolutional Neural Networks (CNNs) supervisadas para aprender y reconocer características y patrones importantes en imágenes de personas. Las CNNs son fundamentales para la detección y ubicación de los puntos clave en el cuerpo, tales como articulaciones, extremidades, manos, cabeza y torso en cada imagen o fotograma. Estos puntos clave dan información para entender la postura y pose de una persona [2].



Figura 2.5: Extracción de la pose en *OpenPose*.

*OpenPose* se basa en dos componentes principales: los mapas de confianza de puntos clave y los Part Affinity Fields (PAF). Los mapas de confianza son mapas de calor que indican la probabilidad de presencia de un punto clave en una ubicación de la imagen. Para cada punto clave, se genera un mapa de confianza separado. Cada píxel en

el mapa de confianza tiene un valor que representa la probabilidad de presencia del punto clave en esa ubicación. Los puntos clave con valores de confianza más altos son considerados más probables. Por otro lado, los PAF son campos vectoriales que contienen información sobre la orientación y la dirección entre dos puntos claves, lo que permite identificar cómo están conectadas las diferentes partes del cuerpo [2]. El uso de estos campos vectoriales brinda ventajas como la detección de múltiples personas en una sola imagen y la capacidad de detectar partes del cuerpo incluso en presencia de obstáculos y pérdida de calidad visual [1]. En la Figura 2.6 se aprecia el uso de los mapas de confianza para la detección de partes y los PAF para la asociación de las mismas.

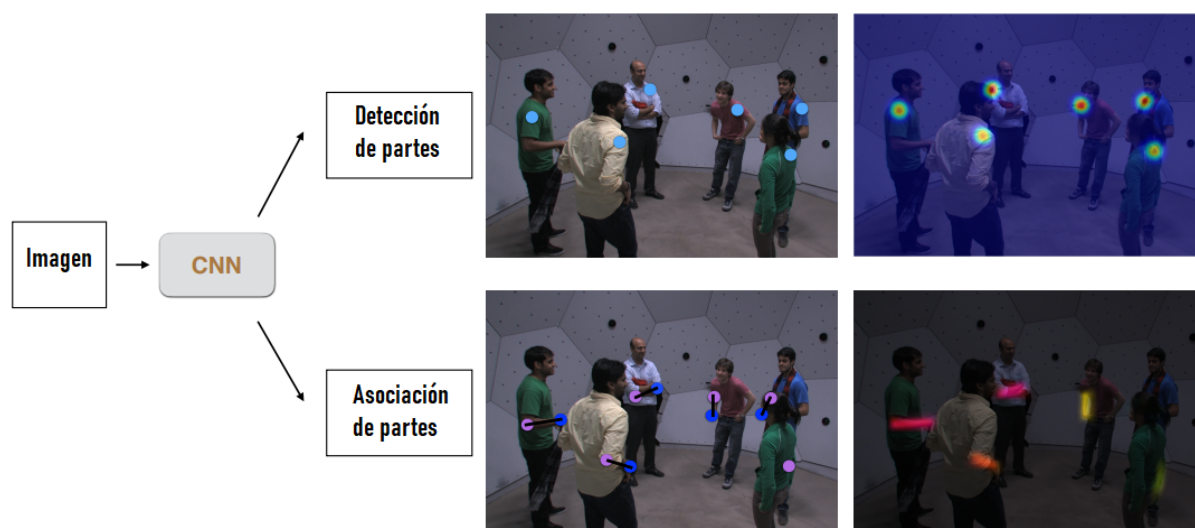


Figura 2.6: Esquema de funcionamiento de *OpenPose*. Fuente: [1]

La Figura 2.7 muestra la arquitectura detrás del funcionamiento de *OpenPose*. El modelo recibe como entrada una imagen en color de tamaño  $h \times w$  y produce, como salida, una matriz de matrices que consiste en los mapas de confianza de puntos clave y los PAF. Los conceptos de capa convolutiva y de *pooling* se explican en la Sección 2.5.1.1

Para comenzar, se generan mapas de características utilizando las primeras 10 capas del modelo VGG-19. Luego, se utiliza una CNN de múltiples etapas con 2 ramas, donde la primera rama predice un conjunto de mapas de confianza  $S$  de las ubicaciones de las partes del cuerpo, un mapa por cada parte. Después, la segunda rama predice

un conjunto de PAF en 2D  $L$ , que codifican el grado de asociación entre los puntos clave. Se tiene un PAF por cada par de puntos clave. Finalmente, los mapas de confianza y los PAF se analizan para obtener los puntos clave en 2D de todas las personas en la imagen. Para esto se traza una línea que une los dos puntos que componen un par cualquiera y se Verifica si los vectores que conforman al PAF de dicho par tienen la misma dirección en los puntos cercanos a la línea. Si la dirección coincide en cierta medida, entonces se considera como un par válido. Los puntos clave del cuerpo se conectan para formar esqueletos sobrepuestos en la imagen o secuencia de video de entrada.

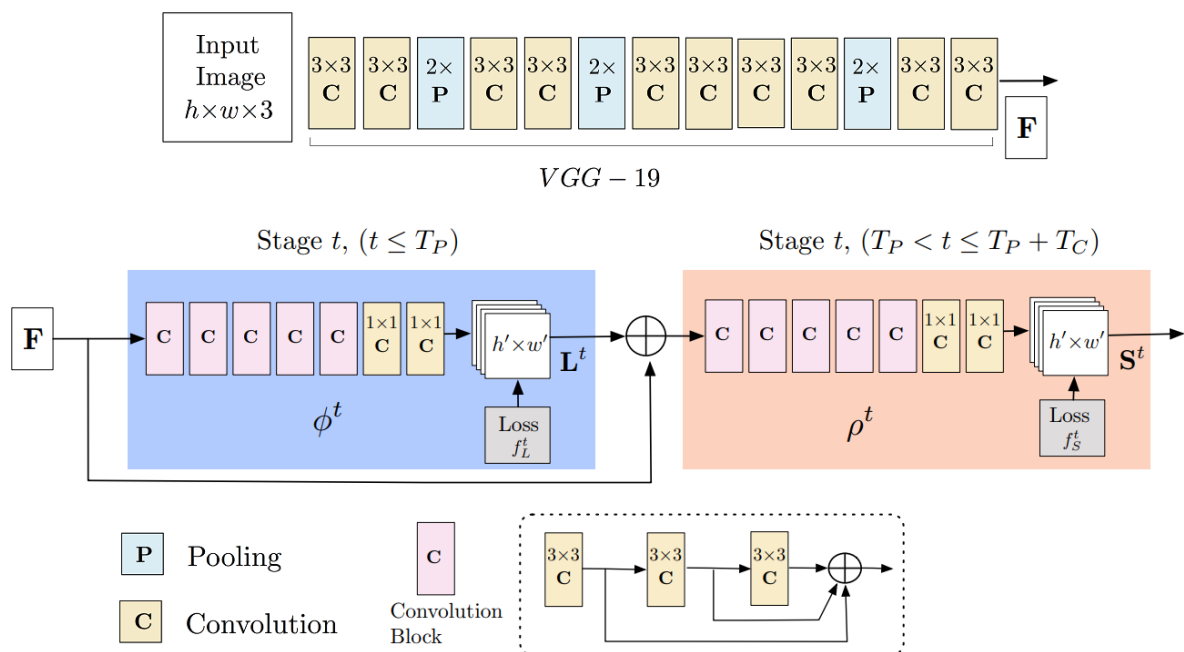


Figura 2.7: Arquitectura de *OpenPose*. Fuente: [1, 2]

Los sistemas de estimación de pose resuelven una variedad de problemas relacionados con el seguimiento e identificación de personas. Sin embargo, uno de los desafíos asociados con estos sistemas es el tiempo de procesamiento computacional elevado que requieren [40]. Por otro lado, *OpenPose* produce modelos con alta exactitud y que no consumen recursos computacionales en exceso [2]. Esta herramienta se ha utilizado en varios trabajos relacionados con el movimiento humano como en la detección de marcha [40, 41] así como en investigaciones en el ámbito deportivo, incluyendo disciplinas como *baseball* [14], *basketball* [42] y atletismo [17].



*OpenPose* es capaz de detectar hasta 135 puntos de referencia y ofrece 4 formatos de salida diferentes, los cuales se detallan en la Tabla 2.1. Estos formatos permiten identificar las partes generales del cuerpo, así como los puntos clave del rostro y las manos.

Tabla 2.1: Formatos de salida de *OpenPose*

<b>Formato de salida</b>	<b>Número de puntos clave</b>
Formato de salida de pose (BODY_25)	25
Formato de salida de pose (COCO)	18
Formato de salida de rostros	70
Formato de salida de manos	21

La Tabla 2.2 muestra los nombres de las partes del cuerpo que se pueden identificar con el modelo BODY\_25 con su respectivo ID. Así mismo, la Figura 2.8 muestra la asignación de las partes del cuerpo del modelo en cuestión.

Tabla 2.2: Orden de mapeo de partes del cuerpo para el modelo *BODY25*

<b>Parte del cuerpo</b>	<b>ID</b>	<b>Parte del cuerpo</b>	<b>ID</b>
Nariz	0	Rodilla izquierda	13
Cuello	1	Tobillo izquierdo	14
Hombro derecho	2	Ojo derecho	15
Codo derecho	3	Ojo izquierdo	16
Muñeca derecha	4	Oreja derecha	17
Hombro izquierdo	5	Oreja izquierda	18
Codo izquierdo	6	Dedo gordo pie izquierdo	19
Muñeca izquierda	7	Dedo pequeño pie izquierdo	20
Cadera media	8	Talón izquierdo	21
Cadera derecha	9	Dedo gordo pie derecho	22
Rodilla derecha	10	Dedo pequeño pie derecho	23
Tobillo derecho	11	Talón derecho	24
Cadera izquierda	12		

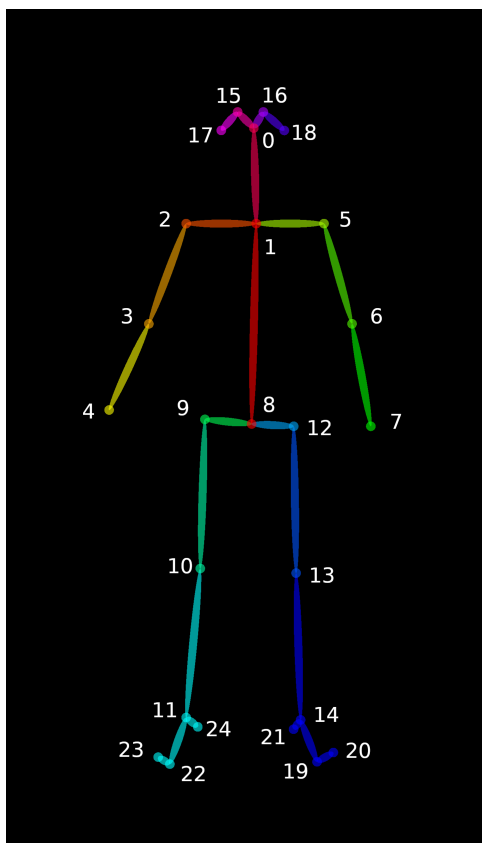


Figura 2.8: Formato de salida de pose (BODY\_25). Fuente: [3]

Siguiendo un proceso similar al realizado, en [43], se extraen las coordenadas horizontales y verticales de los puntos clave detectados en cada fotograma de un video proporcionado al sistema. Sin embargo, es importante tener en cuenta que la precisión de la posición de algunos puntos clave puede no ser ideal. Para mitigar el impacto de los errores de detección, los datos obtenidos en este proceso son sometidos a un tratamiento o preprocesamiento.

## 2.4. Tratamiento de datos

Cuando se trabaja con modelos de Machine Learning (ML) o Deep Learning (DL) es importante verificar que los datos que se ingresan al sistema sean datos de calidad. Este proceso implica transformar los datos recopilados en un nuevo conjunto de información que pueda ser utilizado como entrada para los modelos de ML. Este proceso de transformación de datos incluye varias etapas como reducción del impacto de datos ruidosos o inconsistencias, normalizar los datos para asegurar que se encuentren

en un rango adecuado, seleccionar las características más importantes para el modelo y eliminar datos irrelevantes [44]. Además de obtener datos de calidad, también es posible realizar otras transformaciones en los datos, como lo es la generación de imágenes y de variables espacio temporales, a partir de datos de entrada.

### 2.4.1. Preparación de datos

Sistemas como *OpenPose* han demostrado su eficiencia en la extracción y estimación de la postura humana [45]. Sin embargo, no están exentos de errores, lo que puede generar una fuente de datos poco confiable al momento de proseguir con una aplicación o trabajo determinado. De esta manera, para avanzar en las siguientes etapas de cualquier trabajo que conste de una recolección de datos, es importante atravesar un proceso de depuración y recolección de datos de calidad. Existen varias complicaciones principales asociadas con los datos, como diferencias en las dimensiones, falta de datos o espacios vacíos, presencia de ruido y valores atípicos. Estos problemas pueden abordarse mediante técnicas de normalización, interpolación y suavizado [37].

#### 2.4.1.1. Normalización

Cuando se captura el movimiento de una persona en video, su tamaño relativo experimenta variaciones. Por ejemplo, a medida que un individuo se desplaza hacia la cámara, su figura cambia de tamaño, como consecuencia las medidas de longitud son muy variables. En [46] se utiliza (2.1) para caracterizar el esqueleto sin importar la posición en la que se encuentre.

$$j'_{t,i} = \frac{(x_{t,i} - x_t^{cm}, y_{t,i} - y_t^{cm})}{h_t(j_t^N, j_t^{tm})} \quad (2.1)$$

En (2.1),  $t$  es el número de imagen o fotograma en la secuencia. El término  $j'_{t,i} = (x_{t,i}, y_{t,i})(i = 0, \dots, 24)$  se refiere al valor escalado que se desea calcular de las coordenadas  $x, y$  de una parte específica del cuerpo, donde las coordenadas,  $cm$  hace referencia al punto medio de la cadera. Mientras que  $j_t$  se refiere al valor entregado

por el sistema *OpenPose*.  $h_t$  representa la distancia euclidiana entre  $N$  que hace referencia a las coordenadas de la parte del cuerpo de la nariz, y  $tm$  que es la posición media entre el tobillo derecho y el tobillo izquierdo.

#### 2.4.1.2. Interpolación

Los datos nulos o inexistentes se completan mediante un proceso de interpolación. La técnica de interpolación consiste en estimar valores desconocidos de una variable utilizando los valores conocidos en puntos cercanos [47]. La forma más común y práctica de interpolación es la lineal. La fórmula para la interpolación lineal se muestra en (2.2).

$$y = y_1 + \frac{(y_2 - y_1)}{(x_2 - x_1)}(x - x_1) \quad (2.2)$$

En (2.2),  $y$  es el valor interpolado que se busca calcular, y  $x$  es el valor de entrada para el cual se desea obtener dicha interpolación. Además, se utilizan los valores  $y_1, y_2, x_1, x_2$  que se consideran como datos conocidos necesarios de las variables  $y$  y  $x$ .

#### 2.4.1.3. Técnicas de suavizado de datos

Para abordar problemas como el ruido, errores de detección y valores atípicos en los datos, se puede recurrir al suavizado de datos. El suavizado tiene como objetivo reducir las interferencias y fluctuaciones no deseadas, eliminando el ruido y resaltando los patrones y tendencias significativas en un conjunto de datos [48]. Una técnica comúnmente utilizada es el uso de una ventana de *Hanning*. En términos matemáticos, la ventana de *Hanning* se define como una función coseno al cuadrado con un periodo que abarca toda la longitud de la señal analizada [49, 50].

$$w[n] = 0.5 - 0.5 \cos\left(\frac{2\pi n}{N-1}\right) \quad 0 \leq n \leq N-1 \quad (2.3)$$

Siendo  $N$  la longitud de la ventana, la cual define la cantidad total de muestras in-

cluidas en la función de *Hanning*. Esta ventana se aplica a los datos para atenuar los extremos y reducir las discontinuidades en la transición de una señal [51].

La función de la ventana de *Hanning* asigna un valor de 0 en los bordes de la ventana, con un peso menor, y aumenta sinusoidalmente hacia los valores del centro de la ventana [50], otorgándoles un mayor peso que a los valores del borde. Su forma se asemeja a una curva de campana, como se muestra en la Figura 2.9.

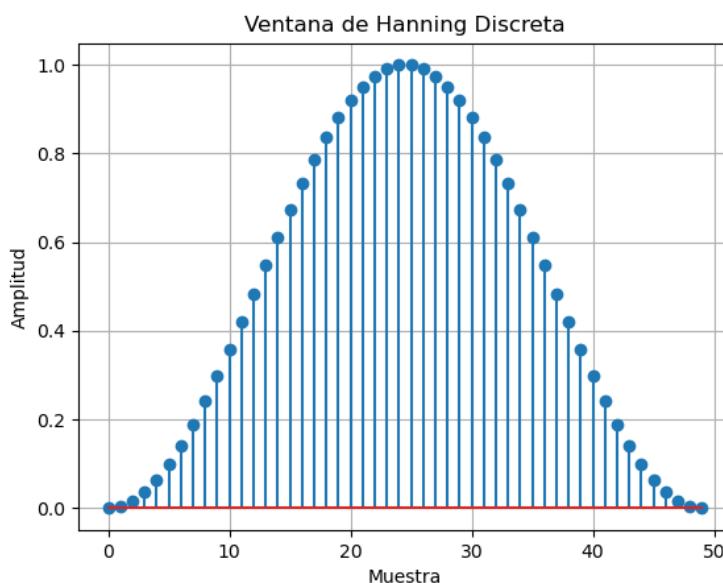


Figura 2.9: Ejemplo de una ventana de *Hanning* con  $N=50$

En el dominio del tiempo, el suavizado mediante la ventana de *Hanning* consiste en convolucionar la señal original con la la ventana. Es decir, implica desplazar la ventana a lo largo de la señal original y calcular el producto punto a punto entre los valores de la ventana y los de la señal original. Esto se expresa en (2.4).

$$y[n] = x[n] * w[n] = \sum_{k=-\infty}^{\infty} x[k] \cdot w[n - k] \quad (2.4)$$

En donde  $x[n]$  es la señal original,  $w[n]$  la ventana de *Hanning* y  $y[n]$  la señal suavizada. Además, la ventana se puede ajustar a diferentes tamaños según las necesidades, y el tamaño determina la cantidad de suavizado aplicado a la señal.

Una vez aplicadas las correcciones y transformaciones necesarias, se obtiene un conjunto de datos de mayor calidad en comparación con los datos iniciales, como se

puede observar en la Figura 2.10. Esto proporciona una entrada mejorada para los sistemas de ML que utilizan estos datos.

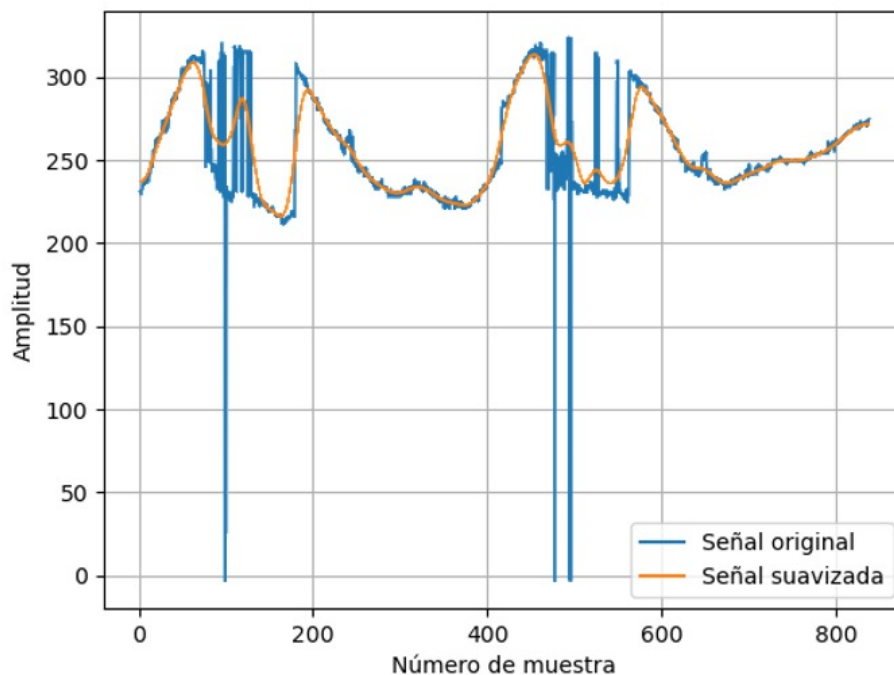


Figura 2.10: Curva antes y después del suavizado.

#### 2.4.2. Imagen de energía de la marcha (GEI)

El Gait Energy Image (GEI) consiste en una imagen promedio de las siluetas binarias generadas por un individuo durante un ciclo de marcha. Se obtiene a partir de una secuencia de imágenes de la marcha, lo que reduce el impacto de los errores presentes en una sola imagen en el resultado global [52]. Esta imagen resultante codifica tanto información estática como dinámica, capturando características fisiológicas y espacio-temporales durante la marcha, como la forma y apariencia del cuerpo, la longitud del paso, la cadencia y las variaciones en frecuencia y fase [53]. El GEI se utiliza ampliamente para la identificación individual y la clasificación de patrones de caminata anormales. Además, la imagen GEI se puede utilizar como entrada en sistemas de aprendizaje automático, como las CNN.

La imagen GEI se calcula a partir de una secuencia temporal de  $N$  imágenes utilizando la fórmula propuesta en [54], que se muestra en (2.5).

$$GEI(x, y) = \frac{1}{N} \sum_{t=1}^N B_t(x, y) \quad (2.5)$$

En (2.5),  $B_t(x, y)$  representa el valor del pixel en las coordenadas  $(x, y)$  de una imagen específica  $B_t$  donde  $t$  indica el número de fotograma. La Figura 2.11 ilustra una secuencia de imágenes de siluetas binarias y muestra el resultado obtenido al aplicar la técnica del GEI.



Figura 2.11: *Gait Energy Image* GEI. Fuente: [4].

Se puede utilizar una variante especial del GEI, el Skeleton Gait Energy Image (SGEI), en el cual en lugar de emplear la silueta completa de la persona, se utiliza únicamente la silueta de su esqueleto. Para lograr esto, es necesario extraer o estimar el esqueleto del individuo en cuestión, lo cual puede realizarse mediante sistemas como *OpenPose* [52] (ver Figura 2.12).

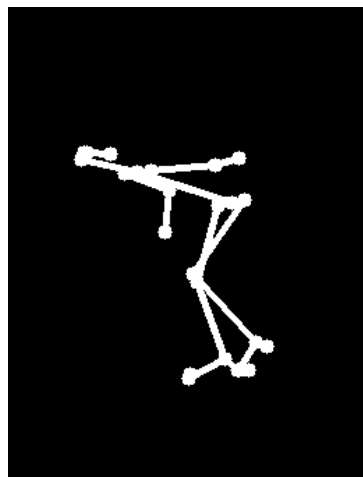


Figura 2.12: *Skeleton Gait Energy Image* (SGEI).

### 2.4.2.1. Distancia entre dos puntos en dos dimensiones

La distancia Euclidiana entre dos puntos  $P = (x_1, y_1)$  y  $Q = (x_2, y_2)$  en un plano de dos dimensiones se calcula utilizando el teorema de Pitágoras [55]. La fórmula correspondiente se muestra en (2.6). Esta fórmula permite determinar la distancia directa entre dos puntos en el plano cartesiano.

$$d(P, Q) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (2.6)$$

### 2.4.2.2. Ángulo formado por tres puntos

Para encontrar el ángulo formado por tres puntos en un plano cartesiano, se utilizan la ley del Coseno. Según esta ley, en un triángulo cualquiera con lados  $a$ ,  $b$  y  $c$ , vértices  $A$ ,  $B$  y  $C$ , y ángulo opuesto a  $c$  denotado por  $\theta$ , se cumple [56] en (2.7).

$$c^2 = a^2 + b^2 - 2ab \cos \theta \quad (2.7)$$

Con base en esta fórmula se obtiene:

$$\theta = \cos^{-1} \left( \frac{a^2 + b^2 - c^2}{2ab} \right) \quad (2.8)$$

Donde  $\theta$  es el ángulo buscado, y  $a$ ,  $b$  y  $c$  son las longitudes de los lados del triángulo formado por los puntos  $A$ ,  $B$  y  $C$ , respectivamente. Para calcular la longitud de los lados se utiliza (2.6), de esta forma  $a = d(A, B)$ ,  $b = d(B, C)$  y  $c = d(A, C)$ .

## 2.5. Algoritmos de clasificación

Los algoritmos de clasificación de aprendizaje supervisado desempeñan un papel fundamental en la categorización de objetos, datos o eventos en diversas clases o categorías. Son especialmente útiles cuando se trabaja con grandes volúmenes de datos que pertenecen a múltiples grupos o categorías, y se busca una separación precisa entre ellos. Estos algoritmos aprenden a identificar patrones y características



en un conjunto de datos de entrenamiento, y luego aplican ese conocimiento para clasificar nuevos datos en las categorías correspondientes. En esencia, analizan los patrones de un conjunto de datos de entrada y los relacionan con una variable de salida.

En la actualidad, los sistemas de clasificación basados en ML son ampliamente utilizados para implementar algoritmos de clasificación. Estos sistemas tienen la capacidad de aprender de forma autónoma a partir de un conjunto de datos de entrenamiento, y luego aplicar ese conocimiento para clasificar nuevos datos en áreas, como el diagnóstico de enfermedades, el análisis de datos multimedia, la categorización y el filtrado de documentos, así como el análisis de redes sociales, entre otros [57].

### 2.5.1. Redes neuronales convolucionales (CNN)

Las CNN son un sistema de aprendizaje automático ampliamente utilizado en el procesamiento de señales e imágenes. Las CNNs utilizan un algoritmo de optimización denominado descenso del gradiente estocástico (Stochastic Gradient Descent (SGD)) para entrenar y optimizar sus parámetros. Se ajustan los pesos de la red para que las salidas se aproximen lo máximo posible a las etiquetas de los datos de entrenamiento [9].

Su arquitectura se compone de varias capas que trabajan en conjunto para extraer características y realizar la clasificación o predicción de los datos de entrada. Las CNN utilizan capas convolucionales que se encargan de aplicar convoluciones entre filtros (*kernels*) y la imagen de entrada. Cada filtro es una matriz que se desplaza sobre la imagen aplicando convolución en cada posición. Cada kernel tiene la capacidad de extraer características específicas de las imágenes de entrada como bordes, texturas o formas [58].

Además, se utilizan capas de agrupamiento o *pooling* para reducir el espacio de la dimensión de los mapas de características. Una forma común de *pooling* es el *max pooling*, que selecciona el valor máximo dentro de una región de la imagen, reduciendo su tamaño y hacer la red más profunda. Las capas densas están completamente conectadas y se utilizan para la clasificación final o predicción. Las funciones de ac-

tivación también desempeñan un papel crucial, ya que se aplican a la salida de cada neurona para introducir no linealidades y mejorar la capacidad de la red para aprender características complejas [59].

Asimismo, las capas de *Flatten* y de regularización son capas adicionales que usan después de las capas de convolución y *pooling*. La capa *Flatten* convierte los mapas de características en un vector unidimensional para poder conectar las capas densas. Mientras que La capa de *dropout* o capa de regularización desactiva de forma aleatoria un porcentaje de las neuronas durante el entrenamiento para evitar el sobreajuste y mejorar la generalización del modelo.

Es importante destacar que en la arquitectura de CNN varía según la aplicación y del modelo que se esté utilizando. Es posible usar múltiples capas como de convolución, *pooling* o adicionales, para mejorar el rendimiento. Además, los tamaños de los filtros (*kernels*) también pueden variar según los requisitos del problema. El tamaño de estos filtros comúnmente son pequeños, pero varían dependiendo de la profundidad y complejidad del modelo [59].

#### **2.5.1.1. Modelo *Visual Geometry Group* (VGG19)**

Visual Geometry Group (VGG19) es una red neuronal convolucional profunda ampliamente utilizada en la clasificación de imágenes. Fue desarrollada por un grupo de investigadores del VGG19 en la Universidad de Oxford en 2014 y ha sido ampliamente empleada en desafíos de reconocimiento de imágenes como el *ImageNet Large Scale Visual Recognition Challenge* [60]. La arquitectura VGG19 consta de 19 capas, incluyendo 16 capas convolucionales y 3 capas completamente conectadas. Las capas convolucionales extraen características de las imágenes de entrada, mientras que las capas completamente conectadas se utilizan para la clasificación. VGG19 es conocida por su alta precisión en la clasificación de imágenes y su arquitectura simple y comprensible la convierte en un punto de partida común para el desarrollo de nuevas arquitecturas [9].

VGG19 hace uso de múltiples filtros de 3x3 en cada una de sus 16 capas convolucionales para extraer características de las imágenes de entrada. Las capas restantes,

que constan de tres capas completamente conectadas, se encargan de la clasificación. Los filtros son matrices cuyas dimensiones dependen del número de canales de la imagen de entrada. En imágenes en escala de grises, los filtros son matrices bidimensionales, mientras que en imágenes RGB se requieren filtros tridimensionales [58]. La salida resultante de aplicar un filtro convolucional se llama *feature maps* o mapas de características.

Las capas convolucionales se agrupan en cinco grupos, y cada grupo está seguido por una capa de *maxpooling* o agrupación máxima. El modelo se alimenta con imágenes de 224x224 píxeles y genera una etiqueta que describe el objeto presente en la imagen [61]. La estructura del modelo se muestra en la Tabla 2.3 y la Figura 2.13.

Tabla 2.3: Modelo 19. Fuente: [9]

Capa	Número de kernels	Tipo de Capa	Dimensión
1	64	Convolucional	3 x 3
2	64	Convolucional	3 x 3
3	-	Pooling	2 x 2
4	128	Convolucional	3 x 3
5	128	Convolucional	3 x 3
6	-	Pooling	2 x 2
7	256	Convolucional	3 x 3
8	256	Convolucional	3 x 3
9	256	Convolucional	3 x 3
10	256	Convolucional	3 x 3
11	-	Pooling	2 x 2
12	512	Convolucional	3 x 3
13	512	Convolucional	3 x 3
14	512	Convolucional	3 x 3
15	512	Convolucional	3 x 3
16	-	Pooling	2 x 2
17	512	Convolucional	3 x 3
18	512	Convolucional	3 x 3
19	512	Convolucional	3 x 3
20	512	Convolucional	3 x 3
21	-	Pooling	2 x 2
22	4096	Completamente conectadas	-
23	4096	Completamente conectadas	-
24	1000	Softmax	-

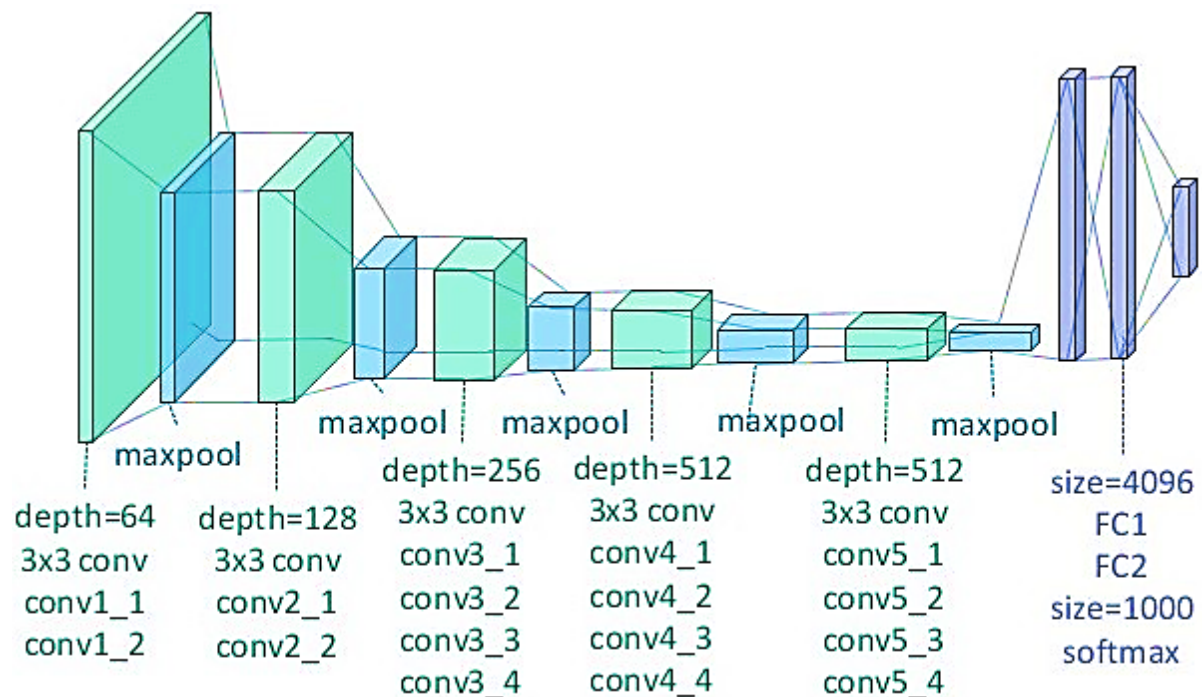


Figura 2.13: Modelo 19. Fuente: [5]

## 2.5.2. Máquinas de vectores de soporte (SVM)

Las (Support Vector Machines (SVMs)) son modelos ampliamente utilizados en ML para la clasificación. Estos modelos supervisados se emplean para determinar la clase o categoría a la que pertenece un dato de entrada, incluso en situaciones con dos o más clases. Las SVM logran esto mediante la creación de un hiperplano que separa y distingue las clases como se visualiza en la Figura 2.14. Su objetivo consiste en encontrar el hiperplano que mejor separa las clases en un espacio de alta dimensionalidad, maximizando el margen o espacio existente entre ellas. El margen se define como la distancia entre el hiperplano separador y los puntos más cercanos a este hiperplano en cada clase. En otras palabras, el margen es la distancia entre las dos líneas paralelas que pasan por los puntos más cercanos a la línea de separación. Así, el propósito de las SVM es encontrar el hiperplano que maximiza dicho margen [6].

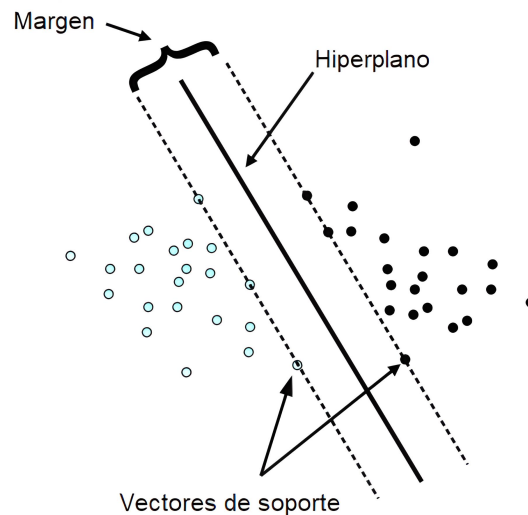


Figura 2.14: Clasificación con SVM. Fuente: [6].

Para encontrar el hiperplano, se abordan dos tipos, el problema primal y el dual. La función del hiperplano para el problema primal de una clase en específico se puede expresar como:

$$f_a(x) = w_a^T x + b_a \quad (2.9)$$

En (2.9),  $w_a$  representa el vector de pesos del clasificador de la clase  $a$ .  $x$  es el vector de características de la muestra de prueba, mientras que  $b_a$  es el sesgo del hiperplano. Se busca encontrar valores óptimos para  $w$  y  $b$  de cada clase, y estos permitan separar las muestras del entrenamiento en las clases deseadas.

El problema dual es una forma alternativa del problema primal mediante el uso de la optimización y sin la necesidad de conocer detalladamente los vectores  $w$  y  $b$  y se puede expresar como:

$$f(x) = \sum_{i=1}^N \alpha_i y_i K(x, x_i) + b \quad (2.10)$$

Donde,  $f(x)$  es la función de decisión,  $N$  corresponde el número de datos de entrenamiento, mientras que  $\alpha_i$  son los coeficientes de *Lagrange* siendo  $y_i$  la etiqueta de clase de la  $i$ -ésima instancia y  $x_i$  el  $i$ -ésimo vector de características.  $K(x, x_i)$  representa una función de kernel que calcula similitudes entre el vector de características

$x$  que se desea predecir su clase y el valor  $i$ -ésimo conocido  $x_1$ .

Una vez encontrado los multiplicadores de Lagrange, los cuales representan la importancia de cada muestra en el proceso de clasificación, se procede a calcular los pesos  $w$  y  $b$  utilizando la siguiente fórmula:

$$w = \sum_{i=1}^N \alpha_i y_i x_i \quad (2.11)$$

El problema dual permite encontrar el hiperplano óptimo para clasificar las muestras en un espacio de características, lo que facilita el uso de funciones de kernel, que permite resolver problemas de clasificación no lineales.

Las SVM se utilizan para dividir el dominio de los datos, y existen dos categorías: lineales y no lineales. En las SVM lineales, se realiza un mapeo de los datos y luego se divide el dominio. Por otro lado, en las no lineales, se utiliza la función de *kernel* para mapear los datos a un espacio de características, es decir, transforma los datos de entrada para permitir su separación mediante un hiperplano en un espacio de mayor dimensión [7] (ver Figura 2.15). Esto facilita encontrar un hiperplano que pueda separar las clases, incluso en casos de patrones no lineales, como secciones circulares [62].

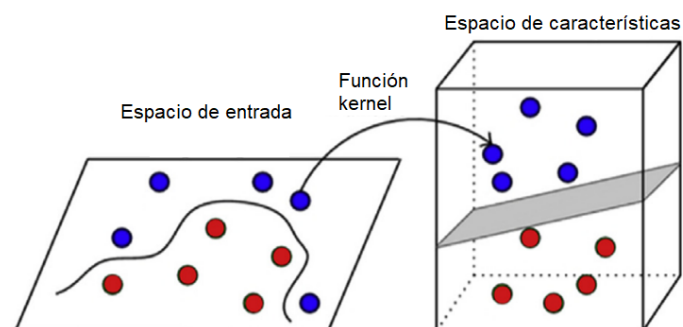


Figura 2.15: Uso del *kernel trick* para introducir datos en un espacio de características de alta dimensionalidad. Fuente: [7].

Una vez que se encuentra el hiperplano óptimo, las SVM se utilizan para clasificar nuevos datos en relación con su posición respecto al hiperplano identificado. Las SVM tienen diversas aplicaciones, como la clasificación de imágenes, detección de spam, clasificación de documentos y predicción de actividad biológica [7, 62].

## 2.6. Evaluación de algoritmos de clasificación

La evaluación de la exactitud de un modelo de clasificación en la predicción y clasificación de datos se realiza mediante pruebas con conjuntos de datos no vistos previamente. Esto permite determinar si el modelo se adapta correctamente a nuevos casos y si es útil en situaciones reales.

Existen diversas medidas para evaluar la capacidad predictiva del modelo en relación a datos nuevos. Algunos de estos métodos incluyen la matriz de confusión, la curva Receiver Operating Characteristic (ROC), la precisión y la exhaustividad, el *F-score*, la validación cruzada, el *bootstrap* y el análisis de curvas de aprendizaje. A continuación se describen brevemente algunos de estos enfoques de evaluación.

### 2.6.1. Matriz de confusión

Una matriz de confusión, también conocida como matriz de error, es una tabla que muestra el desempeño de un método de clasificación al asignar instancias a diferentes clases. La tabla está compuesta por filas y columnas, donde cada fila representa los valores reales de cada clase y cada columna representa las instancias predichas por el modelo [63]. La matriz de confusión proporciona información sobre qué clases están siendo confundidas por el modelo y muestra el recuento de valores predichos y reales. Al analizar la matriz de confusión, se obtiene una visión detallada de las predicciones realizadas por el modelo.

En el caso de la clasificación binaria, se tienen los valores mostrados en la Tabla 2.4 [64]:

- Verdadero positivo (*True Positives*, TP): las instancias que son positivas y son correctamente predichas como positivas.
- Falso positivo (*False Positives*, FP): las instancias que son negativas, pero son incorrectamente predichas como positivas.
- Verdadero negativo (*True Negatives*, TN): las instancias que son negativas y son correctamente predichas como negativas.

- Falso negativo (*False Negatives*, FN): las instancias que son positivas, pero son incorrectamente predichas como negativas.

Tabla 2.4: Matriz de confusión para una clasificación binaria

		Valores Predichos	
		Negativos	Positivos
Valores actuales	Negativos	TN	FP
	Positivos	FN	TP

### 2.6.2. Curva ROC

Una curva ROC es un gráfico como se visualiza en 2.16 que muestra el rendimiento de un sistema de clasificación binario para diferentes valores del umbral de clasificación. La curva ROC representa la tasa de verdaderos positivos o True Positive Rate (TPR) en función de la tasa de falsos positivos o False Positive Rate (FPR) [8]. Los valores de TPR y FPR se especifican en el intervalo de 0 a 1. El valor de TPR también se denomina como Sensibilidad y el valor de 1-FPR se denomina Especificidad. Estos valores pueden ser determinados matemáticamente para graficar la curva ROC [65].

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2.12)$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} \quad (2.13)$$

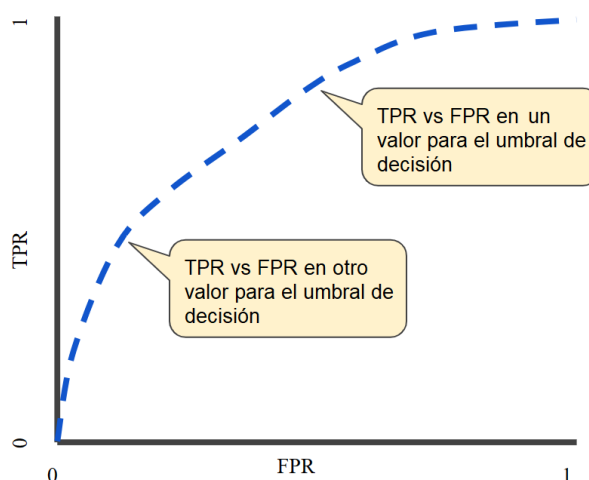


Figura 2.16: Curva ROC. Fuente:[8].



El Area Under the ROC Curve (AUC) es un valor utilizado para medir el rendimiento de un modelo de clasificación y se calcula a partir de la curva ROC. El AUC en 2.17 representa el área bajo la curva ROC y cuanto mayor sea este valor, mejor será el rendimiento del modelo [64]. La curva ROC se encuentra delimitada entre los valores 0 y 1 en ambas dimensiones, por lo que el valor del área también estará limitado entre 0 y 1.

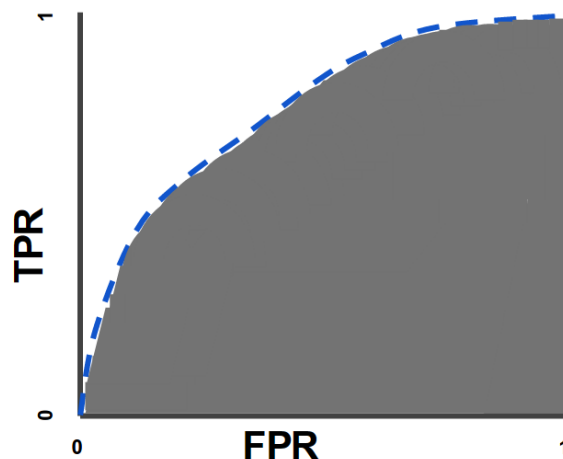


Figura 2.17: Área bajo la curva ROC. Fuente: [8].

### 2.6.3. Otras métricas de evaluación

Existen diversas métricas que se pueden derivar de los datos reflejados en la matriz de confusión. La elección de estas métricas suele estar condicionada por el tipo de problema en cuestión, y por lo general, se pueden identificar las medidas de evaluación a partir de la descripción del problema, a diferencia de la selección del algoritmo utilizado para abordarlo [66].

#### 2.6.3.1. Exactitud

La exactitud se refiere al porcentaje de predicciones correctas en relación al total de predicciones realizadas.

$$\text{Exactitud} = \frac{TP + TN}{TP + FP + TN + FN} \quad (2.14)$$

### 2.6.3.2. Precisión

La precisión es el porcentaje de valores positivos que se clasificaron correctamente.

$$\text{Precisión} = \frac{TP}{TP + FP} \quad (2.15)$$

### 2.6.3.3. Sensibilidad

La sensibilidad es la proporción de verdaderos positivos con relación al total de casos positivos en el conjunto de datos. Las Ecuaciones (2.16) y (2.12) hacen referencia a la misma métrica.

$$\text{Sensibilidad} = \frac{TP}{TP + FN} \quad (2.16)$$

### 2.6.3.4. F1-score

Este valor es la media armónica de la precisión y la sensibilidad. Este valor solo será alto si tanto la precisión como la sensibilidad son altas.

$$\text{F1-score} = 2 \cdot \frac{\text{Precisión} \cdot \text{Sensibilidad}}{\text{Precisión} + \text{Sensibilidad}} \quad (2.17)$$

### Capítulo 3 Trabajos relacionados

Este capítulo presenta el estado de arte y los trabajos relacionados con el trabajo de titulación actual, abordando temas como el reconocimiento de personas, la identificación de acciones en videos y la clasificación de movimientos en el ámbito deportivo. Se inicia con un análisis general de los procesos de análisis del movimiento humano y luego se profundiza en herramientas y procesos implementados en sistemas de aprendizaje automático en el deporte. El enfoque principal se centra en la clasificación precisa de técnicas y movimientos específicos en un deporte determinado. El objetivo es evaluar la viabilidad de utilizar técnicas de aprendizaje automático para obtener información detallada sobre la técnica de los deportistas, con el propósito de mejorar sus futuros entrenamientos.

El ser humano se involucra en una amplia variedad de deportes con distintos objetivos. Sin embargo, para alcanzar los más altos estándares, los deportistas se someten a un arduo entrenamiento y requieren poseer las condiciones físicas y psicológicas necesarias, así como dominar las habilidades y técnicas específicas de cada deporte. Por otro lado, en la actualidad, los seres humanos tienen la capacidad de capturar información en diversos formatos, siendo los archivos de video uno de los más populares. Estos archivos se generan diariamente en grandes cantidades y contienen información de diversos tipos, convirtiéndose en una fuente de datos de gran importancia y consideración. Ante esta creciente avalancha de datos y los avances tecnológicos en áreas como el Machine Learning (ML), el Deep Learning (DL), la Computer Vision (CV) y el procesamiento de imágenes, han surgido nuevos campos que están ganando terreno en la investigación científica. Uno de estos campos es el Human Action Recognition (HAR) que ha experimentado un desarrollo y aplicaciones en áreas como la generación de resúmenes de video basados en contenido, la interacción humano-computadora, la educación, la atención médica, la vigilancia mediante video, la detección de actividades anómalas y el deporte [67]. En el contexto de este trabajo, el deporte se presenta como el principal punto de interés, ya que los entrenamientos deportivos siempre buscan la innovación para superar barreras establecidas y mejorar el rendimiento en una disciplina deportiva.

### 3.1. HAR y aprendizaje automático en el ámbito deportivo

En el artículo [68], se examina el uso de diversos métodos ML y DL en investigaciones relacionadas con deportes específicos. Se presentan dos enfoques para implementar el reconocimiento de actividad humana (HAR): el enfoque tradicional de ML con extracción manual de características y el enfoque de DL con ingeniería automática de características. Se destaca la importancia de disponer de conjuntos de datos etiquetados para el entrenamiento supervisado y se mencionan desafíos como la complejidad de las acciones, grupos de muestras desbalanceados, oclusión y movimientos rápidos de cámara en las bases de datos disponibles, lo cual dificulta la aplicación de métodos de ML y DL. Se observa que el 17,27 % de los trabajos analizados emplea el método Support Vector Machine (SVM), mientras que las Convolutional Neural Network (CNN) se utilizan en un 51,81 % de los casos.

Así mismo, en [69] se presenta una revisión de técnicas para el reconocimiento y predicción de acciones en videos. Se identifican dos etapas fundamentales en el proceso de reconocimiento de acciones: la representación de acción y la clasificación de acción. Respecto a la clasificación de acciones, se destaca la importancia de considerar la información de las relaciones geométricas entre las diferentes partes del cuerpo. Algunos de los métodos utilizados incluyen SVM, K Nearest Neighbors (KNN) y K-means clustering. El estudio también aborda protocolos de evaluación, como la matriz de confusión, para medir la precisión del reconocimiento en bases de datos pequeñas mediante SVM y el método de entrenamiento *leave-one-out*. Además, se emplean métricas como la sensibilidad, el Area Under the ROC Curve (AUC) y la precisión promedio, la cual se aproxima al área bajo la curva de precisión-sensibilidad.

En relación a deportes específicos, [70] examina el reconocimiento de acciones en el balonmano utilizando una red basada en Long Short-Term Memory (LSTM). El sistema LSTM se alimenta con una secuencia de características extraídas de videos mediante la red *InceptionV3* con los pesos pre-entrenados de *ImageNet*. En el reconocimiento de acciones, se aplican dos enfoques: el primero abarca 9 categorías de acciones, mientras que el segundo contempla 11 categorías. Con el primer enfoque, se logra una exactitud de validación del 70.94 %, mientras que con el segundo se

alcanza una exactitud de validación del 70.55 %.

Por otro lado, en [71] se emplean sensores ubicados en las raquetas y técnicas de ML para el reconocimiento de acciones en el bádminton. En dicho trabajo, se recopilan datos correspondientes a diez acciones principales de este deporte. Los resultados obtenidos muestran que una CNN (AlexNet) logra una exactitud de reconocimiento del 90.24 %, mientras que LSTM alcanza el 85.18 % y SVM solo llega al 22.97 %. Sin embargo, mediante la incorporación de un bloque de extracción de características adaptativas, la CNN logra un 98.65 % de exactitud en el reconocimiento. El bloque adaptativo realiza un proceso similar al Principal Component Analysis (PCA), permitiendo mejorar significativamente los resultados obtenidos.

Además, se han realizado otros análisis interesantes en el campo, como el presentado en [72], que propone un marco para la clasificación de imágenes deportivas en función de su entorno y contexto. Para este estudio, la base de datos es creada a partir de videos de YouTube y es dividida en conjuntos de entrenamiento y prueba en una proporción de 70-30. Luego, el modelo de reconocimiento de imágenes *InceptionV3* es empleado para extraer características relevantes de las imágenes. Posteriormente, se utilizan diversos métodos de ML para llevar a cabo la clasificación de las imágenes. Se implementa una CNN, SVM, KNN y Random Forest (RF), logrando valores de exactitud del 96.64 %, 94.77 %, 92.58 % y 90.10 %, respectivamente, para la clasificación en seis categorías de deportes posibles. Se emplea la matriz de confusión como medida para evaluar la efectividad de los clasificadores utilizados.

### 3.2. OpenPose como herramienta de HAR

En relación al reconocimiento de personas mediante CV, existen diferentes enfoques. Uno de ellos es el enfoque de arriba hacia abajo, que consiste en aislar a las personas en una imagen y luego identificar sus acciones. Sin embargo, este enfoque puede requerir un gran esfuerzo computacional. Como alternativa, se utiliza software como *OpenPose* para la detección, el cual extrae el esqueleto de una o varias personas utilizando puntos clave como articulaciones o extremidades [2]. Esta herramienta aplica HAR a nivel de pose. En [73] se realizan investigaciones para explorar la aplicabilidad

de *OpenPose* en un entorno de vigilancia de video. El estudio propone un enfoque que extrae características de la pose y las utiliza como datos de entrada para una red basada en LSTM, con el objetivo de clasificar 17 tipos de acciones. Se aplican tres enfoques utilizando una base de datos existente: utilizar los datos sin modificar, aplicar *Data Augmentation* a los datos y utilizar validación cruzada. Estos enfoques arrojan valores de exactitud del 92.73 %, 95.58 % y 96.44 %, respectivamente. El valor de validación cruzada se obtuvo entrenando la red con la base de datos preexistente y probándola con una base de datos creada específicamente para el estudio.

En [74], se presenta un sistema que extrae el esqueleto en 2D a partir de un conjunto de datos utilizando *OpenPose*. Las características principales de los datos del esqueleto son extraídas para así obviar la etapa de extracción de características presente al ingresar la información a la arquitectura que combina CNN y LSTM para el reconocimiento de acciones. En este estudio se emplean dos bases de datos: la primera conformada por 18 categorías de acciones y un total de 14298 muestras, mientras que la segunda incluye 12 categorías y 1800 muestras. Se utiliza una partición de entrenamiento y prueba con una proporción de 80-20. Los resultados muestran una precisión del 94.4 % en la primera base de datos y del 91.67 % en la segunda. Para evaluar el desempeño, se utiliza la matriz de confusión y las métricas de precisión, la sensibilidad y F1-score.

En un estudio similar, descrito en [75], se emplea *OpenPose* para capturar las poses de un individuo formando un conjunto de 32 fotogramas que se utilizan como entrada en una red LSTM con el objetivo de reconocer actividades humanas en tiempo real. El trabajo se enfoca en diferenciar 6 categorías de acciones, logrando una exactitud de clasificación del 87.12 %.

Además, en [76], se utiliza también *OpenPose* para obtener el esqueleto de una persona y crear vectores de características de movimiento a partir de secuencias de video. Para identificar las actividades relacionadas con los vectores de características obtenidos, se emplea una Recurrent Neural Network (RNN) con celdas LSTM. La base de datos utilizada contiene 11 categorías de acciones. El modelo LSTM logra una exactitud de clasificación del 92.4 %. Para hacer una comparación, se implementa un modelo RF que alcanza el 78.5 % de exactitud. Además, se utilizan matrices de confu-

sión, así como las métricas de precisión y sensibilidad como medidas de evaluación.

En el ámbito deportivo, se han realizado estudios que aplican *OpenPose* para diversas finalidades. Por ejemplo, en [14], se propone un método para medir el nivel de los *swings* de un bateador de béisbol utilizando *OpenPose*. En este estudio, se utilizan las coordenadas de las extremidades del bateador para calcular la distancia de la cadera y los ángulos de las extremidades. Estas variables se evalúan utilizando reglas personalizadas basadas en la investigación y la experiencia de entrenadores profesionales. El estudio muestra una correlación positiva entre la corrección de las acciones y el desempeño de los bateadores.

Asimismo, en [42], se emplea *OpenPose* para crear un modelo de análisis de postura en el baloncesto. Se utiliza un modelo de regresión logística y la postura del esqueleto para predecir la probabilidad de encestar un tiro libre. La evaluación de los resultados se da mediante el *accuracy ratio*, que es la proporción de la mejora de rendimiento del modelo en comparación con un modelo aleatorio, en relación con la mejora de rendimiento de un modelo perfecto en comparación con el modelo aleatorio. Se obtiene un valor de *accuracy ratio* de 41 %.

### 3.3. Gait Energy Image (GEI) como un enfoque de CV en diferentes aplicaciones

En el ámbito de la visión por computadora, se exploran diferentes enfoques. Uno de ellos es el análisis de la silueta, que ha dado lugar al desarrollo de herramientas como el GEI. En el estudio [77], se utiliza el GEI, obtenido a partir de ciclos de caminata, para estimar la edad de personas entre 15 y 40 años. Se forma un vector de características y se escala a un subespacio de menor dimensión que luego se utiliza para entrenar un clasificador basado en KNN. Se utiliza la métrica del error absoluto medio para evaluar el rendimiento del modelo, y se obtiene un valor de 3.7 años.

Continuando con este enfoque, en [78] se emplea el GEI para alimentar una CNN profunda con el objetivo de prevenir caídas en personas vulnerables, como adultos mayores o personas con discapacidad. El modelo basado alcanza una exactitud de clasificación del 99.1 % y una tasa de predicción del 98.64 %. Además, se evalúan otras métricas como la especificidad, sensibilidad y precisión de detección para vali-

dar el rendimiento del modelo propuesto. Por otra parte, en [69], en la etapa de representación de acción, se analizan diversos estudios que emplean herramientas como el Motion Energy Image (MEI) y el Motion History Image (MHI) para extraer características y generar vectores de características. Estas herramientas comparten el mismo principio que el GEI, el cual se centra en el análisis del movimiento en secuencias de imágenes.

De manera similar, en el artículo [79] se propone la utilización de imágenes de energía combinando los principios de GEI, MEI y MHI en conjunto con dos arquitecturas de CNN: el modelo de Visual Geometry Group (VGG19) y el modelo Residual Network (ResNet)-18. Estas técnicas se evaluaron en dos bases de datos públicas, obteniendo resultados de exactitud del 96.49 % y 81.25 % para ResNet-18, y del 99.58 % y 96.56 % para VGG19.

### 3.4. Construcción de conjuntos de datos

En el contexto del desarrollo de métodos de detección basados en aprendizaje automático, es importante contar con un conjunto de datos adecuados, considerando la calidad y cantidad de los datos recopilados. En este contexto, existen varios retos y enfoques para crear el conjunto de datos de estudio. En [70] se explica el proceso de captura de videos durante un partido de balonmano utilizando cámaras estacionarias. Sin embargo, los videos resultantes presentan movimientos rápidos de los jugadores y un fondo ruidoso. Para abordar estos problemas, se emplea un enfoque que obtiene una secuencia de características a partir de los fotogramas de un video mediante la red *InceptionV3*, previamente entrenada con los pesos de *ImageNet*. Además, se emplea el optimizador Adaptive Moment Estimation (ADAM) durante un máximo de 100 épocas para mejorar el rendimiento del modelo. Esta secuencia se utiliza como entrada al sistema de clasificación basado en LSTM. Por otro lado, en [69] se analizan varios trabajos de HAR y se aborda diversos problemas asociados a la captura de video, como el movimiento de la cámara, las condiciones de iluminación, los cambios de punto de vista y el fondo dinámico o ruidoso. Además, se enfrenta a la variación intraclase, que se refiere a las diferencias en el movimiento dentro de una misma categoría, ya que cada persona realiza una misma acción de manera ligeramente distinta.



Para superar estos desafíos, el estudio examina diferentes enfoques de extracción de características. Se consideran métodos como la aplicación de MEI y MHI, así como el uso de la ecuación de Poisson para extraer propiedades utilizadas en la representación y clasificación de acciones. También se emplean algoritmos de *OpticalFlow*, que revelan los patrones de movimiento aparente entre dos fotogramas consecutivos. Además, se utiliza información de mapas de profundidad y datos del esqueleto capturado para mejorar la capacidad de reconocimiento. Adicionalmente, el estudio explora el potencial de las redes profundas para la extracción automática de características a partir de videos.

En [80] se parte de tres bases de datos: NATOPS, SBU Kinect y BodyLogin, que contienen información de movimiento de diferentes gestos realizados por sujetos de prueba. Cada una de estas bases de datos proporciona mapas de profundidad, a partir de los cuales se generaron secuencias de siluetas binarias para cada acción realizada. En este trabajo, se plantea el uso de imágenes de energía derivadas de estas siluetas binarias como entrada a sistemas de ML. Una forma común de obtener imágenes de energía es mediante el método del GEI que promedia todos los fotogramas y les asigna la misma relevancia. Sin embargo, este enfoque no aprovecha al máximo la información disponible cuando las acciones clave de un movimiento tienen una duración corta. Para superar esta limitación, el método propuesto en este estudio introduce pesos adaptativos basados en la importancia de cada fotograma. Dicha importancia se mide mediante la correlación entre los fotogramas, lo que resalta las características más relevantes de un movimiento sin importar su duración. De esta manera, se logra un enfoque más efectivo para capturar y utilizar la información clave de los movimientos en las imágenes de energía.

Referente a las bases de datos públicas disponibles, estas no siempre satisfacen las necesidades de un proyecto específico, lo que lleva a varios estudios a crear sus propios conjuntos de datos. En el estudio realizado por [71], se recopilan datos mediante un sensor colocado en una raqueta, registrando 10 acciones diferentes con un promedio de 119 muestras por acción. De manera similar, en el estudio de [81], se utiliza un sensor en una raqueta para distinguir 3 tipos de golpes, recopilando en promedio 70 muestras por tipo, con un total de 1080 muestras de un grupo profesional y 648 de

un grupo de principiantes. Estos estudios cuentan con la supervisión de profesionales para verificar la correcta identificación de las categorías. Además, en el estudio [73] se construye una base de datos con 18 acciones propuestas, capturadas mediante una cámara Red Green Blue (RGB) estática, generando en promedio 40 muestras por acción y aplicando técnicas *Data Augmentation* para obtener un total de 5598 muestras. Por otro lado, en el estudio de [82] se describe un método para construir un conjunto de datos de reconocimiento de acciones en el deporte de balonmano utilizando sistemas de reconocimiento de objetos basados en DL, logrando un rendimiento del 84 % de aciertos en un conjunto de datos reales recopilados de videos de práctica.

### **3.5. Aplicación de técnicas de aprendizaje automático para la clasificación de técnicas y movimientos en deportes**

En el ámbito de la clasificación de movimientos del cuerpo humano, se han llevado a cabo varios trabajos que aplican técnicas de ML, como se describe en los estudios de [37] y [83]. En estos trabajos, se utiliza una CNN y un sistema SVM con datos obtenidos a través de *OpenPose* para clasificar marchas patológicas que afectan las extremidades inferiores. Los resultados obtenidos muestran una exactitud del 98.57 % para la CNN y del 96.87 % para SVM.

Netamente en el ámbito deportivo, se han realizado diversos estudios para clasificar e identificar las técnicas y movimientos utilizados por deportistas en diferentes disciplinas. El estudio [84] describe el uso de una CNN para la clasificación de golpes de tenis, utilizando información temporal representada como gráficos y técnicas de Attention Mechanisms (AT). Se utilizan datos tridimensionales registrados mediante un sistema de captura de movimiento, y se logra una exactitud del 85,54 % al emplear datos solo de la silueta del jugador y 93 % al emplear datos de la silueta completa del jugador junto con la raqueta de tenis. De manera similar, en [85] se presenta una investigación que compara diferentes algoritmos de ML para clasificar 13 golpes de pádel, basándose en datos recopilados mediante una Inertial Measurement Unit (IMU). En este estudio se realiza una evaluación del rendimiento de diversos algoritmos, entre los cuales se incluyen Fully Connected Neural Network (FCNN), CNN de una dimensión, árbol de decisión, KNN y SVM. Los resultados obtenidos muestran

que la FCNN alcanza una exactitud del 92.6 %, la CNN de una dimensión obtiene una exactitud del 93.35 %, el árbol de decisión logra una exactitud del 62.09 %, SVM alcanza una exactitud del 91.85 % y KNN obtiene una exactitud del 84.84 %. El estudio también analiza métricas de precisión, sensibilidad y F1-score además de matrices de confusión.

Continuando con el enfoque de clasificación, se han realizado investigaciones relacionadas con la identificación de técnicas en deportes específicos. Por ejemplo, [86] utiliza un modelo basado en CNN y LSTM para clasificar las técnicas de esquí de fondo (*Cross-country skiing*) con el objetivo de mejorar el rendimiento de los esquiadores profesionales. Se comparan los resultados del modelo propuesto con un sistema KNN. El modelo CNN-LSTM obtiene una exactitud de 95.1 % mientras que el modelo KNN obtiene una exactitud de 78 %. Por otro lado, en [87], se emplea un giroscopio acoplado al tobillo de los patinadores para obtener medidas durante la ejecución de rutinas de movimientos. Se utiliza un análisis cuantitativo y diferentes técnicas de clasificación, como árbol de decisión, RF, KNN y SVM para diferenciar a los patinadores en niveles de principiante, intermedio y profesional. Los métodos logran exactitudes de 91.8 % para el árbol de decisión, 92.8 % para RF, 72.86 % para KNN y 83.4 % para SVM en el grupo de patinadores profesionales. En [88], se logra la clasificación automática de subtécnicas de esquí de fondo mediante el uso de sensores IMU y redes neuronales implementadas en el *toolbox* de MATLAB. Esta metodología alcanza una exactitud del 93,9 %.

Siguiendo esta línea de investigación, en [81] se realiza la clasificación de diferentes tipos de golpes de raqueta en tenis de mesa utilizando una base de datos creada a partir de señales de un sensor. Se implementan tres sistemas: SVM con Radial Basis Function (RBF), una CNN bidimensional y LSTM, obteniendo valores de exactitud de 76.96 %, 75.66 % y 81.30 %, respectivamente. Además, se utilizan métricas como la precisión, sensibilidad y F1-score. Aunque el sistema LSTM demuestra un rendimiento superior a los otros dos métodos en el estudio, los parámetros de SVM resultan más fáciles de configurar.

En relación a las SVM, [89] utiliza este sistema para mejorar la precisión del reconocimiento de objetivos en competiciones deportivas, abordando desafíos como el brillo

desigual y los cambios bruscos que pueden generar errores de reconocimiento. El estudio propone un sistema de visión artificial que emplea un modelo SVM para el reconocimiento de individuos obteniendo valores de al rededor de 95 % de exactitud.

Estas investigaciones demuestran cómo el uso de diferentes modelos de aprendizaje automático, como LSTM, CNN y SVM, permite clasificar y reconocer técnicas y movimientos específicos en deportes, mejorando la precisión y ofreciendo nuevas oportunidades en el campo del análisis deportivo y el rendimiento de los deportistas.

Considerando los estudios previos mencionados, este trabajo se propone utilizar sistemas de clasificación basados en ML para catalogar los diferentes tipos de empuje en patinadores de velocidad a partir de secuencias de video. Lo que se busca es un análisis y mejora del rendimiento de la técnica de los deportistas, bajo la supervisión de entrenadores capacitados, implementando modelos de CNN, SVM y herramientas como *OpenPose* y el GEI. Esta implementación busca contribuir a la formación de nuevos patinadores, y proporcionar herramientas que complementen el entrenamiento y preparación para competencias de alto nivel. En este trabajo de titulación se realizará un análisis comparativo de los sistemas de clasificación basados en ML aplicados a patinadores de velocidad en el contexto del deporte de alto rendimiento.

## Capítulo 4 Metodología

En el presente capítulo se presenta la metodología utilizada para obtener los datos e información necesarios para alimentar los sistemas de clasificación basados en Machine Learning (ML). En la Sección 4.1 se describe el protocolo de adquisición de toma de datos en la pista de patinaje, incluyendo la distribución de los equipos de captura de video, las dimensiones de la zona de grabación y el protocolo seguido para grabar a los patinadores, además de detallar las herramientas de software y hardware utilizadas. La Sección 4.2 aborda el proceso de extracción de coordenadas mediante *OpenPose* y su pre-procesamiento para disminuir discrepancias entre el valor capturado y el valor real. Luego, en la Sección 4.3 se explica cómo se obtienen representaciones de las características biomecánicas del proceso de patinaje a partir de las coordenadas de los puntos clave del cuerpo. Finalmente, se explica el uso de los métodos de clasificación basados en ML en la Sección 4.4.

### 4.1. Protocolo de adquisición de datos

La captura de datos se realizó en el Patinódromo del complejo deportivo de Totoracocha, Cuenca, Ecuador, contando con el permiso de la entrenadora del club Gabriela Cavaliere (Anexo A). Todos los deportistas que participan en este estudio son voluntarios que firman un consentimiento informado (Anexo B, Anexo C). Los participantes cuentan con edades entre 10 y 24 años. Se tiene la participación de 6 hombres y de 13 mujeres. Cada uno de estos deportistas ejecuta uno de los 3 tipos posibles de técnica de empuje al realizar el movimiento de desplazamiento en recta del patinaje.

Se delimita la zona de la pista sobre la cual se captura el video mediante conos de señalización, tomando en consideración el desplazamiento lateral que los deportistas realizan de forma natural y dejando el espacio suficiente para el mismo. Así mismo, se procura mantener un rango que permita capturar toda la información de un ciclo de patinaje en recta como se menciona en [15] sin comprometer la calidad visual de las vistas frontales y sagitales.

#### 4.1.1. Distribución de equipos

Los equipos de adquisición de video se colocan sobre trípodes y se fijan a una altura de 1.26 metros debido a que esta altura proporciona una vista completa de las acciones y movimientos del patinador durante su desplazamiento. La distribución de las cámaras sobre la pista de patinaje, así como la de los conos de señalización, se basaron en experimentos empíricos. Esta distribución se puede apreciar en la Figura 4.1, con las dos cámaras laterales a una separación de 6.88 metros desde de la pista y 11 metros entre ellas. La cámara frontal se coloca a 9.32 metros de la sección delimitada por conos en la que los patinadores ejecutan la técnica de recta. En la Figura 4.1, se excluye la zona de color rojo de la captura de video, sirviendo únicamente como referencia visual para los patinadores.

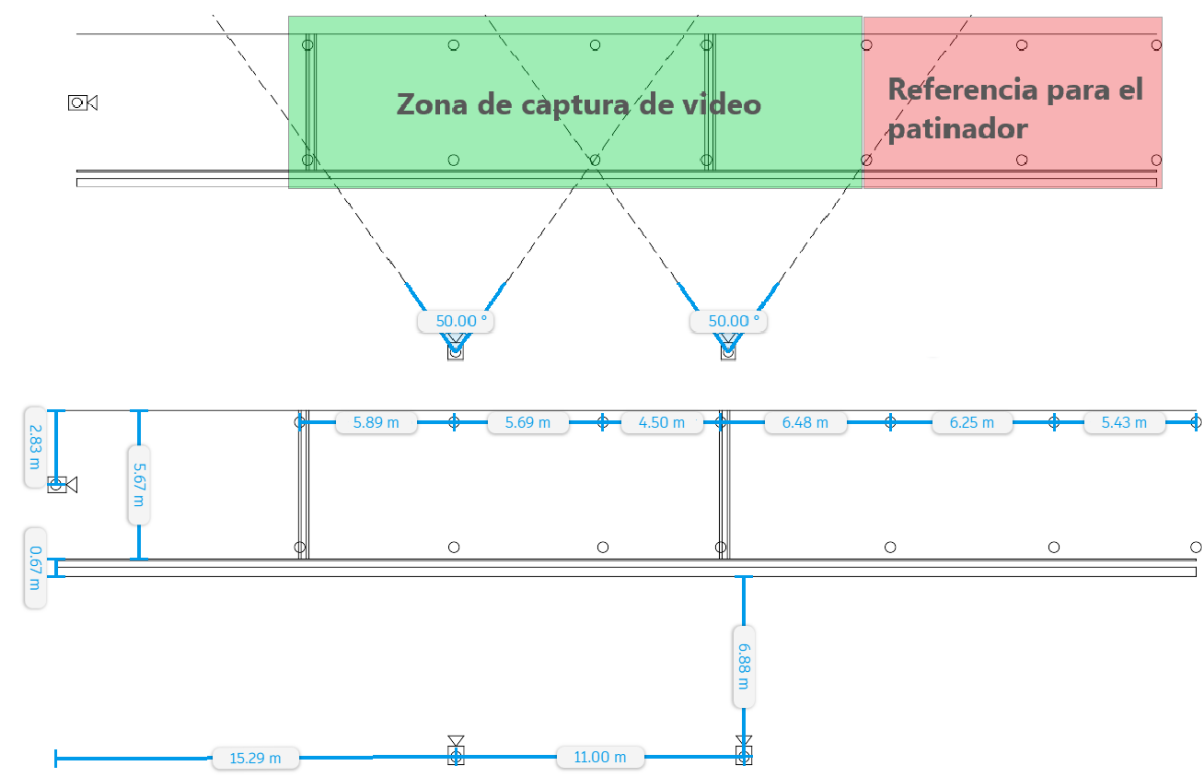


Figura 4.1: Distribución de las cámaras y marcas de referencia.

Durante el proceso de registro del movimiento de los deportistas, se procuró disponer de iluminación natural y ambientar el espacio con música de fondo, con el objetivo de facilitar la posterior sincronización de los videos mediante el software de edición de

video *Filmora* 12.

Al contar con videos que comparten el mismo sonido, el software mencionado permite realizar cortes precisos y unir de manera consistente los diferentes ángulos de grabación. Esta estrategia fue diseñada para garantizar la coherencia temporal entre los distintos planos y obtener una representación fidedigna de los movimientos realizados durante la práctica deportiva.

#### 4.1.2. Descripción de Hardware

Se usa los servicios de supercomputador de Corporación Ecuatoriana para el Desarrollo de la Investigación y la Academia (CEDIA), que brinda acceso a un conjunto de dispositivos interconectados mediante redes de alta velocidad que tienen la capacidad de procesar grandes volúmenes de datos y llevar a cabo diversas operaciones computacionales. La configuración del supercomputador consta de 512 núcleos, 2 TB de Random Access Memory (RAM), 128 TB de almacenamiento y utiliza un sistema operativo basado en Linux [90]. Para la obtención de los datos de pose de los patinadores mediante *OpenPose* se utiliza una partición de 20 GB de Video Random Access Memory (VRAM) de una tarjeta Nvidia A100 habilitada con tecnología Multi-Instance GPU (MIG). Los requerimientos mínimos para la ejecución de este programa son 1.6 GB de VRAM y 2.5 GB de RAM; sin embargo, para obtener mejores resultados y tiempos menores de procesamiento se debe contar con una mayor cantidad de recursos. Esto se debe a que en *OpenPose* se utiliza variables de ajuste, mientras más grande sean estas variables, mayor capacidad de procesamiento y memoria va a requerir el sistema. Por esta razón se opta por el uso de los servicios de supercomputador de CEDIA. Así mismo se utilizan las prestaciones de *Google Colab*, que brinda 12 GB tanto de RAM como de VRAM, para llevar a cabo los procesos de clasificación de datos en *Python*. Este servicio se utiliza debido a la facilidad que ofrece para establecer una comunicación directa con Google Drive, donde se encuentran los datos de entrada para los clasificadores.

#### 4.1.2.1. Descripción de los equipos de captura de video

Se utilizan tres cámaras móviles para capturar el movimiento de los patinadores: dos para la vista sagital y una para la vista frontal. Para la vista frontal, se utiliza la cámara posterior de un iPhone X, que graba video a 60 cuadros por segundo y a una resolución Ultra High Definition (UHD) (3840 píxeles horizontales por 2160 verticales). Para la vista sagital, se emplean dos dispositivos idénticos, el Xiaomi Redmi Note 10 Pro, que graba video a 240 cuadros por segundo y con una resolución High Definition (HD) (1280 píxeles horizontales por 720 verticales). El uso de dispositivos iguales para la vista sagital simplifica las etapas de procesamiento posterior, ya que no es necesario modificar formatos ni resoluciones de video entre ellos. Cada una de estas cámaras se coloca sobre su respectivo trípode para asegurar una altura y distancia uniforme al capturar el movimiento de los patinadores.

#### 4.1.3. Herramientas informáticas

Las herramientas informáticas se utilizan para transformar la información sin procesar en datos que sean capaces de ser suministrados a sistemas de clasificación. Respecto al almacenamiento de datos, se utilizan los servicios de *Google Drive* para almacenar toda la información recopilada en la etapa de captura de video y en etapas posteriores. Para la edición del material visual recopilado se trabaja con el software de edición de video *Filmora 12* que soporta varios formatos de codificación y resolución de video, permite modificar la tasa de cuadros por segundo según sea necesario y cuenta con una opción de sincronización automática basada en audio para diferentes ángulos de vista. Para la extracción de pose, se utiliza *OpenPose* que se instala en el supercomputador de CEDIA, por lo cual el proceso de instalación debe realizarse tomando en cuenta el entorno y comandos de este sistema. El Apéndice A.1 detalla el proceso de instalación. La información obtenida es procesada mediante scripts de Python, con los cuales se obtienen los datos que son introducidos a los sistemas de clasificación.



## 4.2. Pre-procesamiento de datos

Una vez que se hayan registrado todos los videos necesarios con los patinadores involucrados, se contará con una gran cantidad de datos que requerirán ser procesados para extraer información relacionada con el movimiento de las articulaciones y extremidades del cuerpo, con el objetivo de clasificar el tipo de empuje utilizado por el patinador. Este procedimiento, ilustrado en el esquema presente en la Figura 4.2, implica ingresar los videos y obtener resultados en forma de un gráfico Skeleton Gait Energy Image (SGEI) y un conjunto de vectores de características.

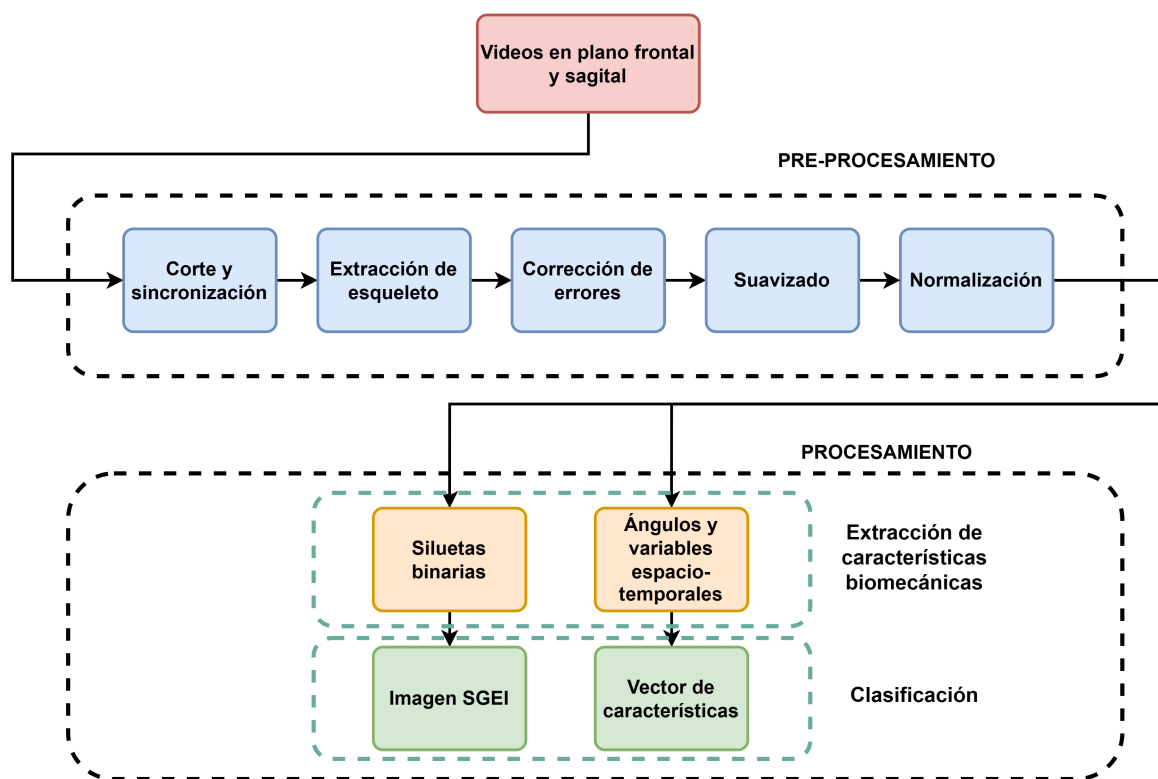


Figura 4.2: Esquema de procesos.

Después de realizar el pre-procesamiento de los datos, se obtienen las coordenadas de las articulaciones y extremidades del cuerpo con valores que se asemejan lo más posible al movimiento real de los deportistas. Estas coordenadas se utilizan en dos procesos: primero, se generan las siluetas binarias del esqueleto en cada fotograma, lo que permite obtener el SGEI posteriormente; segundo, se calculan los ángulos y variables espacio-temporales para obtener un conjunto de vectores de características.

A continuación se detalla cada uno de estos pasos:

#### 4.2.1. Corte y sincronización de videos

Los videos recopilados durante el proceso de toma de datos no tienen la misma duración entre sí. Esto se debe a que se graban tres videos en cada ocasión: dos videos correspondientes a la vista sagital y uno correspondiente a la vista frontal. Además, los videos capturan información antes y después de la ejecución de la técnica de patinaje en línea recta, por lo que es necesario descartar el material innecesario. Esto implica realizar un corte en los videos, eliminando las partes anteriores y posteriores al proceso de patinaje en la zona delimitada.

Luego del corte se procede a sincronizar los videos. El proceso de sincronización consiste en alinear temporalmente dos secuencias de video para que coincidan en tiempo y se reproduzcan simultáneamente de manera coherente. En primer lugar, se realiza la sincronización de las dos vistas sagitales para combinarlas en un solo video correspondiente a esta perspectiva. Luego, se procede a sincronizar este video con el correspondiente a la vista frontal. Para llevar a cabo estos procesos, se utiliza la herramienta de sincronización automática basada en audio del software *Filmora 12*.

#### 4.2.2. Extracción del esqueleto

Para la extracción del esqueleto se utiliza la herramienta de *OpenPose*, que se basa en Convolutional Neural Networks (CNNs) para determinar su información de salida. Esta librería procesa las imágenes de entrada a través de múltiples capas CNN, que aprenden a reconocer y localizar los puntos clave del cuerpo humano.

*OpenPose* se instala en un sistema operativo basado en Linux. Una vez instalada, se llama a la herramienta mediante comandos específicos, los cuales se detallan en el Apéndice A.2. Al introducir un video a *OpenPose*, se obtiene como salida el video con los puntos clave dibujados e interconectados sobre la persona, o una carpeta con archivos JSON que contienen las coordenadas 2D  $(x, y)$  que describen la ubicación de los puntos clave. Dichas coordenadas se miden en píxeles y son relativas a la resolución del video original. Estos formatos de salida se ilustran en las Figuras 4.3 y

## 4.4.

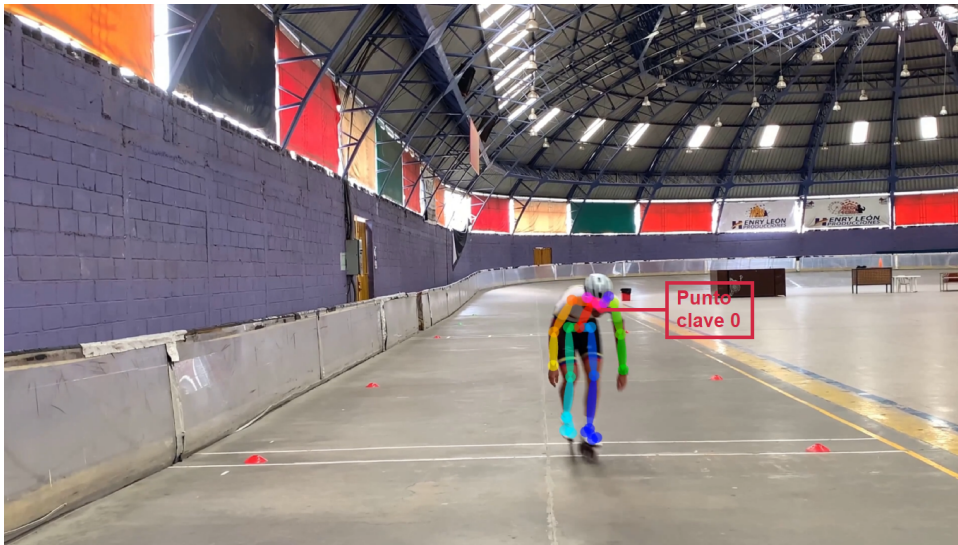


Figura 4.3: Fotograma de un video analizado con el esqueleto sobrepuesto sobre la persona.

```

{"version":1.3,"people":[{"person_id":[-1],"pose_keypoints_2d":
[2378.31,1206.05,0.766567,2340.93,1185.26,0.819407,2261.92,1176.92,0.693459,2189.19,1303.68,0.833058,2187.11,1436.74,0.8811
52,2428.19,1189.37,0.812189,2455.21,1312.05,0.725252,2465.62,1453.35,0.841007,2291.02,1284.99,0.670587,2247.33,1284.99,0.66
5118,2255.66,1482.45,0.841134,2239.09,1640.42,0.750357,2336.74,1285,0.6604,2349.23,1476.21,0.852595,2330.48,1688.24,0.72767
9,2361.66,1187.34,0.754318,2390.77,1187.35,0.685367,2322.19,1166.52,0.573686,2407.41,1162.4,0.0642274,2342.99,1729.8,0.6096
74,2359.6,1723.58,0.703736,2315.94,1700.72,0.561431,2259.85,1704.87,0.614396,2232.82,1694.49,0.589596,2241.11,1644.61,0.568
031],"face_keypoints_2d":[],"hand_left_keypoints_2d":[],"hand_right_keypoints_2d":[],"pose_keypoints_3d":
[],"face_keypoints_3d":[],"hand_left_keypoints_3d":[],"hand_right_keypoints_3d":[]}]}}

```

**Coordenadas**

**2378.31,1206.05,0.766567**      **Coordenada y valor de confiabilidad para el punto clave 0**

X            Y            c

Figura 4.4: Contenido del archivo JSON para el fotograma analizado.

Cada archivo JSON corresponde a un fotograma del video original, y dentro de cada uno se encuentra un diccionario que contiene las coordenadas de cada punto clave identificado, así como el valor de confiabilidad asociado a su detección. En la Figura 4.4 se aprecia el contenido de uno de estos archivos JSON. De esta manera, se forma un conjunto de datos a partir de estos archivos, los cuales se dividen según la persona capturada en el video y el número de videos por persona.

### 4.2.3. Corrección de errores

En esta etapa del pre-procesamiento de datos, el objetivo es corregir los errores en las coordenadas obtenidas de los archivos JSON. Dentro del conjunto de datos, se identifican tres tipos de errores.

**Tipo 1:** Se produce cuando en un determinado fotograma, *OpenPose* no logra identificar un punto clave específico. En tales casos, *OpenPose* asigna automáticamente el valor de  $(0, 0)$  como coordenada para representar la ubicación desconocida del punto clave.

**Tipo 2:** *OpenPose* identifica de forma imprecisa o incorrecta una coordenada específica. Esto puede ocurrir en secciones del video con efectos borrosos debido a la velocidad del movimiento del patinador, problemas de iluminación o sombras.

**Tipo 3:** El tercer tipo de error se relaciona con la identificación intercambiada de partes del cuerpo. En este caso, las coordenadas tienen valores correctos, pero se intercambian entre la parte izquierda y derecha del cuerpo. Este problema se presenta en la vista sagital y afecta principalmente a las extremidades de las piernas. Dado que *OpenPose* utiliza una red neuronal para obtener mapas de confianza para cada del cuerpo, si la persona está en una posición donde la identificación resulta ambigua, *OpenPose* puede llegar a intercambiar la identificación de la parte izquierda con la parte derecha, es por esto que este error se presenta solo en la vista sagital, donde en determinados fotogramas no se puede distinguir de manera adecuada si una pierna está delante de la otra en relación a la cámara.

Para realizar la corrección de errores se expresan los datos de las coordenadas como series temporales. Este enfoque se mantiene en los procesos de suavizado y normalización. La Figura 4.5 muestra el proceso de obtención de las series temporales. En esta Figura, el valor de  $a$  corresponde al identificador del punto clave, mientras que  $b$  indica si se refiere a la coordenada en  $x$  o  $y$ .

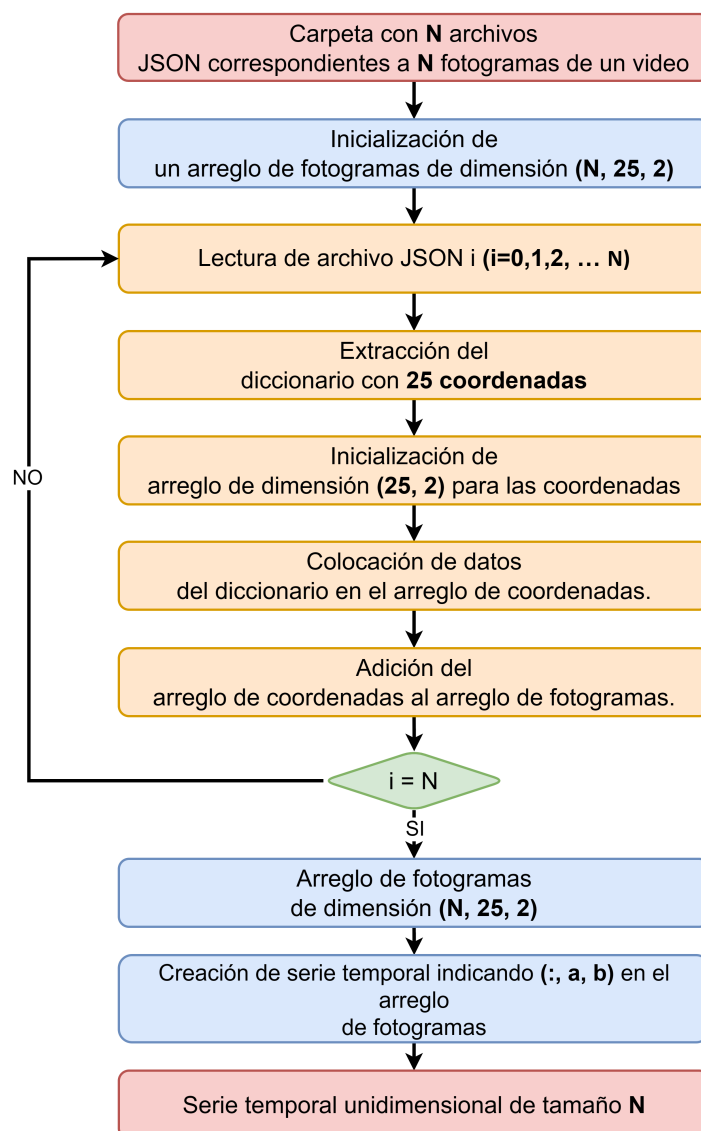


Figura 4.5: Diagrama de flujo para la obtención de la serie temporal.

Para cada punto clave, se obtiene una serie temporal tanto para la coordenada  $x$  como para la coordenada  $y$ . Al ingresar los datos en formato JSON, se genera un conjunto de series temporales con un tamaño igual al número de fotogramas del video analizado. La Figura 4.6 muestra dos series temporales correspondientes a la coordenada  $x$  de los tobillos izquierdo y derecho de un video particular.

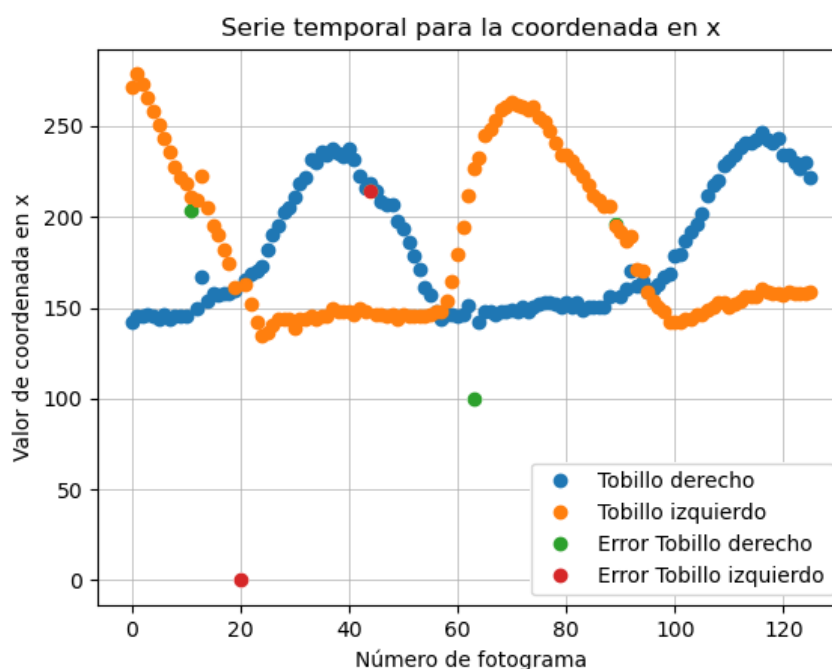


Figura 4.6: Series temporales para las coordenadas en  $x$  de los puntos de los tobillos, Vista Sagital.

La Figura 4.6 muestra la presencia de varios puntos erróneos para los puntos del tobillo. El proceso de corrección de errores se lleva a cabo de forma individual para cada serie temporal. Para ilustrar dicho proceso, se considera la coordenada  $x$  del tobillo derecho como ejemplo. En la Figura 4.7, se resaltan y se detallan los errores correspondientes a dicha serie temporal, donde se identifican los tres tipos de errores especificados.

Una vez obtenida la serie temporal para su análisis, es necesario interpolar los puntos que la componen. Esto permitirá aproximar los datos que presentan errores de tipo 1 a sus valores reales. Con el fin de corregir el resto de los errores, es necesario identificar su ubicación en la serie temporal. Para ello, se utiliza un criterio basado en cambios en el valor de la coordenada. Si se detecta que un dato ha experimentado un cambio significativo en comparación con su valor anterior, se marcará la ubicación correspondiente como una zona donde se produce un error. Para determinar si el cambio es significativo y considerarse como error, se establecen los umbrales indicados en (4.1) y (4.2).

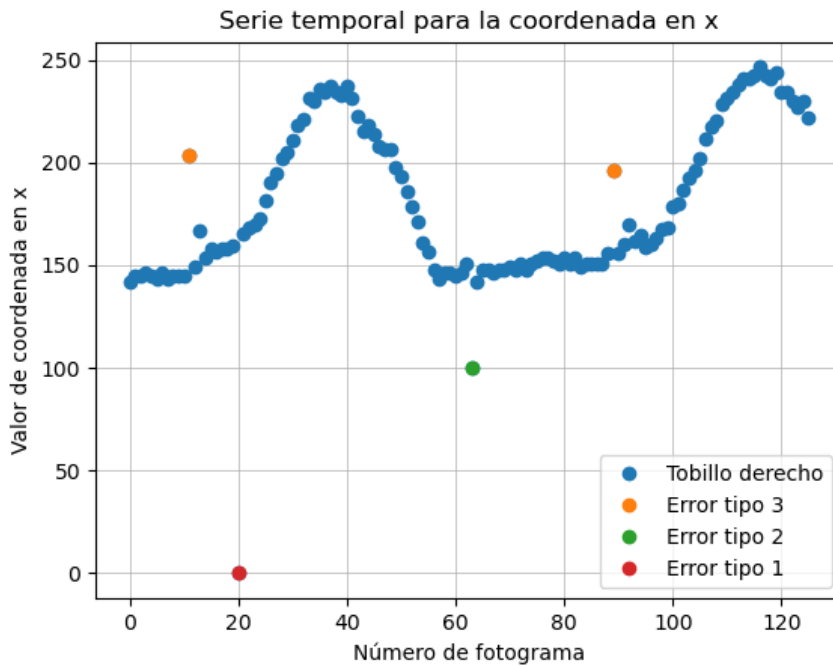


Figura 4.7: Serie temporal para la coordenada en  $x$  del tobillo derecho, Vista Sagital.

$$thr_{up} = \bar{d} + 2 \cdot s_d \quad (4.1)$$

$$thr_{do} = \bar{d} - 2 \cdot s_d \quad (4.2)$$

En (4.1) y (4.2),  $\bar{d}$  representa la media de las diferencias entre valores consecutivos de la serie temporal, mientras que  $s_d$  representa la desviación estándar de estas diferencias. El número de diferencias es igual al tamaño de la serie temporal menos uno. El uso del criterio de 2 desviaciones estándar como indicador de valor anómalo se justifica porque abarca aproximadamente el 95% de los datos en una distribución normal, lo que sugiere que los valores fuera de este rango son estadísticamente inusuales. Se considera que un punto presenta un error si la diferencia entre este y el punto anterior es mayor que  $thr_{up}$  o menor que  $thr_{do}$ .

La pendiente de una recta se define como la relación entre el cambio vertical (diferencia en el eje  $y$ ) y el cambio horizontal (diferencia en el eje  $x$ ) entre dos puntos en la recta. En el caso de la serie temporal, al unir dos puntos consecutivos, las diferencias calculadas representan las pendientes de las rectas formadas. Esto se debe a que el

cambio en el eje  $y$  se obtiene a partir de las diferencias calculadas, mientras que el cambio en el eje  $x$  siempre será igual a 1, ya que representa el número de fotogramas en el video.

En la Figura 4.8, se resaltan con líneas rojas punteadas las ubicaciones donde se producen los saltos identificados como errores. Además, se marcan tanto el punto inicial como el final, con el fin de tener en cuenta todos los datos.

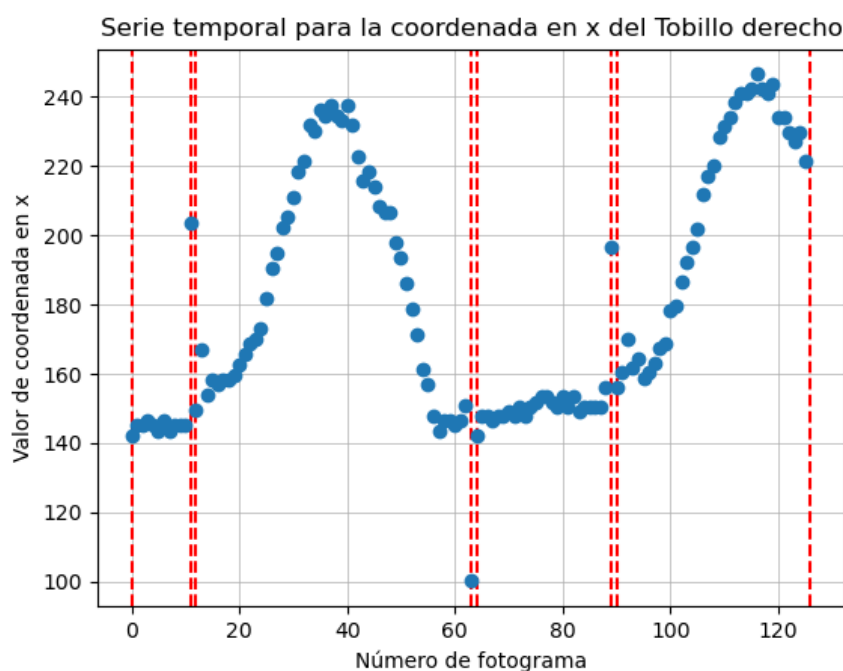


Figura 4.8: Ubicación de los saltos correspondientes a errores en la serie temporal.

El siguiente paso consiste en definir intervalos basados en las ubicaciones de los errores. Se establece un criterio según el cual, si un intervalo consta de 10 o más valores, se considera libre de errores, es decir, que sus valores son confiables. Este criterio se estableció después de analizar varios videos y observar que las ráfagas de errores no excedían los 8 errores consecutivos. Por precaución, se ha definido un valor de 10 como un límite más alto. Por el contrario, si un intervalo contiene menos de 10 valores, se identifica como una zona con errores. En la Figura 4.9, se muestra cómo se divide la serie temporal en intervalos teniendo en cuenta la presencia de errores. Las zonas sin problemas se resaltan en verde, mientras que las zonas con errores se marcan en rojo.



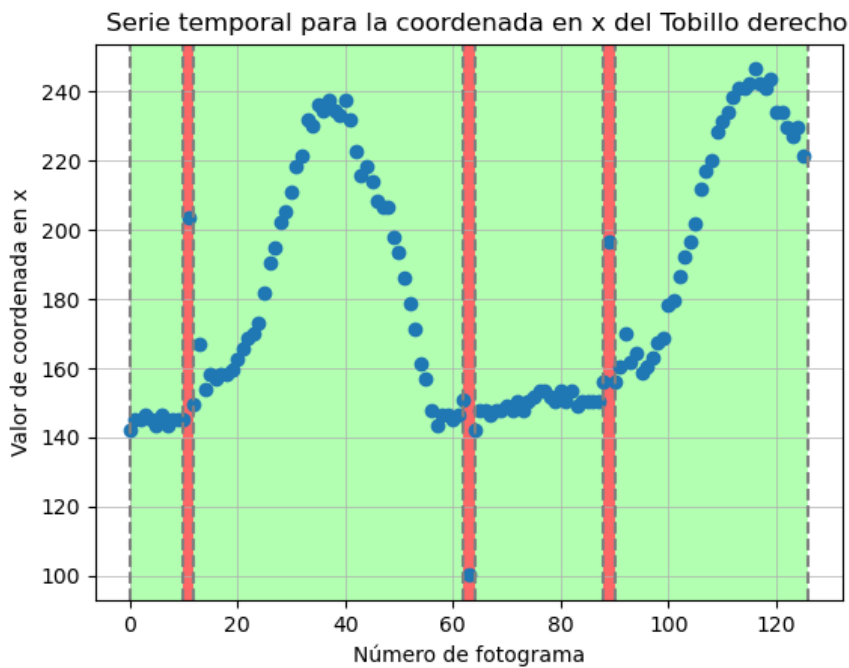


Figura 4.9: Separación en intervalos sin problemas e intervalos con errores.

Una vez definidos los intervalos de error, se procede a aplicar un proceso iterativo de suavizado en las zonas conflictivas. Un diagrama de flujo del proceso se muestra en la Figura 4.10.

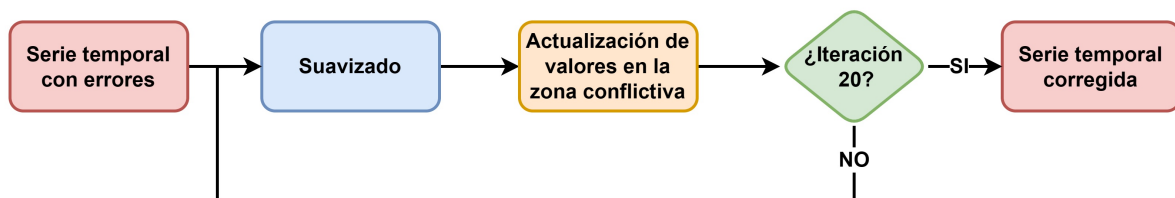


Figura 4.10: Diagrama de flujo para corrección de errores mediante suavizado iterativo.

Primero se aplica un suavizado mediante la convolución de la serie temporal con una ventana discreta de *Hanning* de un tamaño de 10, obteniendo así una nueva serie suavizada. Se selecciona un tamaño de 10 para la ventana, ya que con este tamaño es posible preservar los patrones y tendencias locales presentes en la serie temporal. Después, se procede a reemplazar los valores correspondientes a la zona problemática en la serie original con los nuevos valores obtenidos en la serie temporal

suavizada. Esto asegura que los datos fuera de las zonas conflictivas permanezcan intactos, mientras que los valores erróneos se ajustan gradualmente a la tendencia general de la serie temporal. Este proceso se repite de forma iterativa hasta que los valores erróneos se aproximen lo más posible a la tendencia observada en la serie temporal. En determinados casos, este proceso puede requerir hasta 15 iteraciones para alcanzar la convergencia deseada. Para asegurar un enfoque precautorio, se establece un límite máximo de 20 iteraciones.

En la Figura 4.11, se puede apreciar cómo el valor erróneo correspondiente al fotograma 89 se corrige gradualmente en cada iteración. En este caso particular, se requieren 4 iteraciones para que dicho valor erróneo se ajuste a la tendencia observada en la serie temporal. La Figura 4.12 muestra el proceso para el resto de errores.

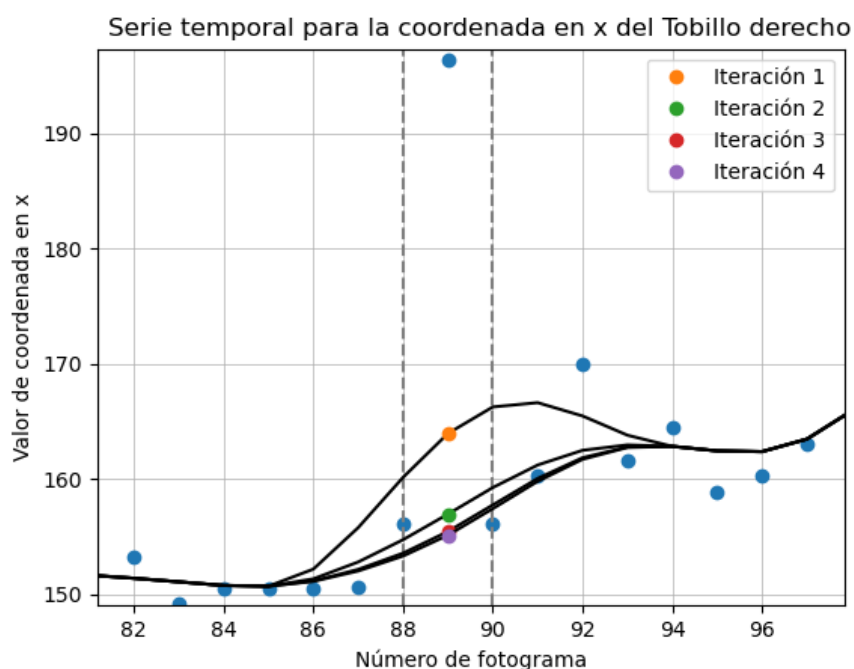


Figura 4.11: Suavizado iterativo para la corrección de errores.

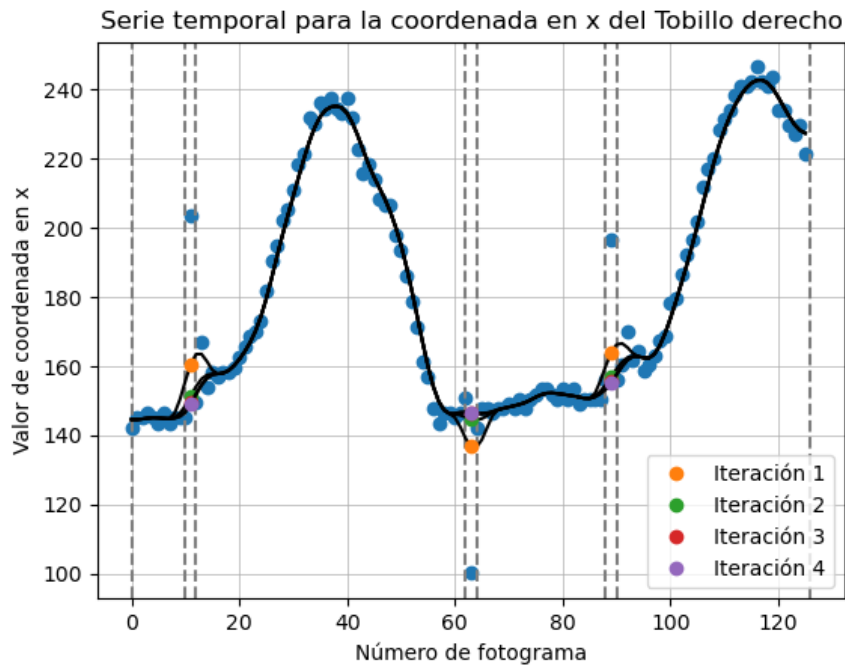


Figura 4.12: Suavizado iterativo para la corrección de errores en la serie temporal.

#### 4.2.4. Suavizado de datos

Luego de realizar el proceso de corrección de errores es importante destacar que aún existen ligeras variaciones en los datos. Con el fin de reducir y eliminar dichas variaciones mínimas, se implementa nuevamente un proceso de suavizado mediante una ventana de *Hanning*, mencionada en la Sección 2.4.1.3 con un tamaño de 10. Este proceso consiste en convolucionar la secuencia temporal con la ventana, permitiendo representar de manera más precisa el comportamiento de la serie. En la Figura 4.13 se presenta la coordenada en  $x$  del tobillo derecho, antes y después de realizar el suavizado.

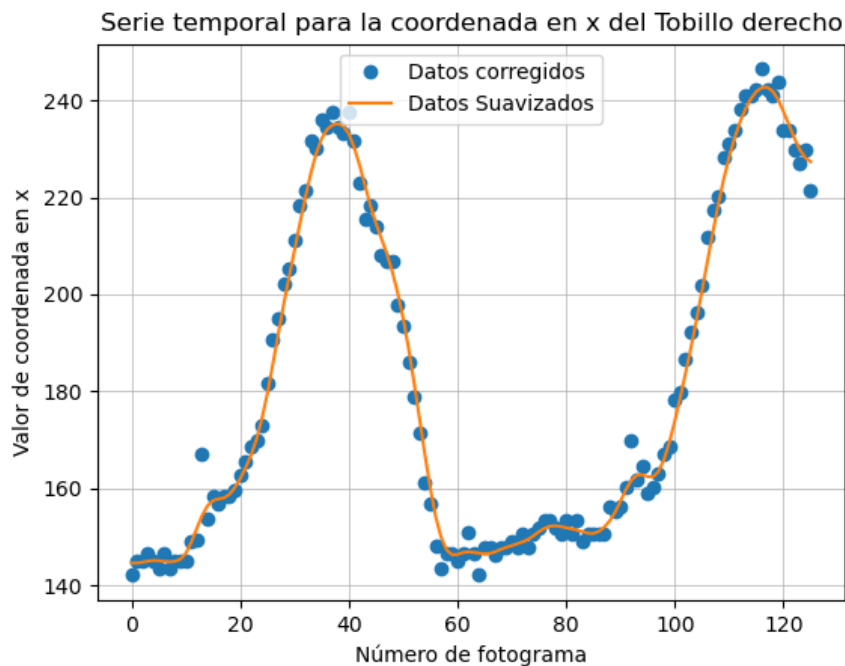
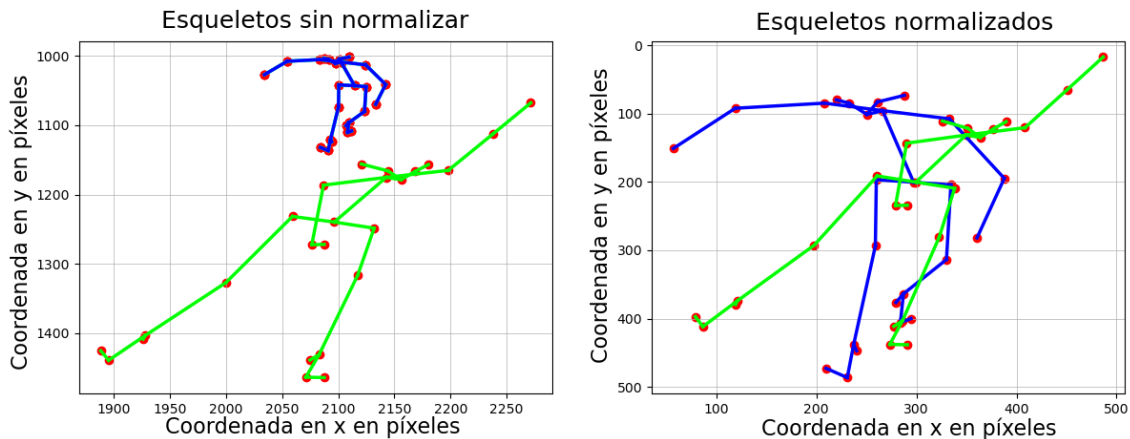


Figura 4.13: Serie temporal con y sin suavizado.

#### 4.2.5. Normalización de datos

En esta etapa se aplica un proceso de normalización a los datos suavizados con el objetivo de asegurar que todos se encuentren en una escala común. Este proceso permite eliminar las disparidades de magnitud entre los datos, facilitando así su comprensión y análisis. Para ilustrar esta sección del pre-procesamiento, se considera la vista frontal, ya que en ella se puede observar de mejor manera el impacto de la variación en el tamaño del deportista a medida que se acerca a la cámara frontal (Ver Figura 4.14(a)). Para realizar la normalización, se sigue el método detallado en la Sección 2.4.1.1.



(a) Datos de un individuo en dos frames diferentes sin normalizar. (b) Datos de un individuo en dos frames diferentes aplicando normalización.

Figura 4.14: Normalización de datos en la vista frontal.

Para realizar la normalización, primero se establece un valor de distancia  $h_t$  constante en relación al movimiento del cuerpo en cada fotograma. Esto se logra utilizando la distancia desde la nariz hasta el punto medio entre los dos tobillos. A continuación, se calculan las posiciones relativas de los puntos clave con respecto al punto central de la cadera en cada fotograma. Estas nuevas posiciones se escalan dividiéndolas por la distancia  $h_t$  correspondiente en cada fotograma. De esta manera, los valores de las posiciones resultantes siempre estarán comprendidos entre -1 y 1, ya que la distancia desde la nariz hasta el punto medio entre los dos tobillos es mayor que cualquier otra distancia entre un punto clave y el punto central de la cadera. Finalmente, para representar adecuadamente las nuevas posiciones en una imagen, se escalan los valores que se encuentran entre -1 y 1 multiplicándolos por un factor de 300.

En la Figura 4.15, se muestra el resultado de aplicar la normalización a la serie temporal de la coordenada  $x$  del tobillo derecho en la vista frontal. En la Figura, la serie temporal suavizada, representada en color azul, se encuentra por encima de valores de 2000 en el eje vertical, el cual representa la posición en píxeles en el video. Al aplicar el proceso de normalización, solo se considera la posición de los puntos clave en relación con el punto central de la cadera. Por ejemplo, si en un fotograma se tiene un valor de 2100 píxeles para la coordenada  $x$  del tobillo derecho y un valor de 2000 píxeles para la coordenada  $x$  del punto central de la cadera, se toma en cuenta solo la separación existente entre las dos coordenadas, es decir 100 píxeles. Como re-

sultado, los valores normalizados están "centrados" alrededor de 0, ya que ese valor ahora corresponde a la posición de la cadera central. La normalización garantiza que todos los datos se encuentren en la misma escala (-1, 1). Además, se observa cómo la forma general de la serie se ajusta después de la normalización, lo que contribuye a una representación más consistente de los puntos.

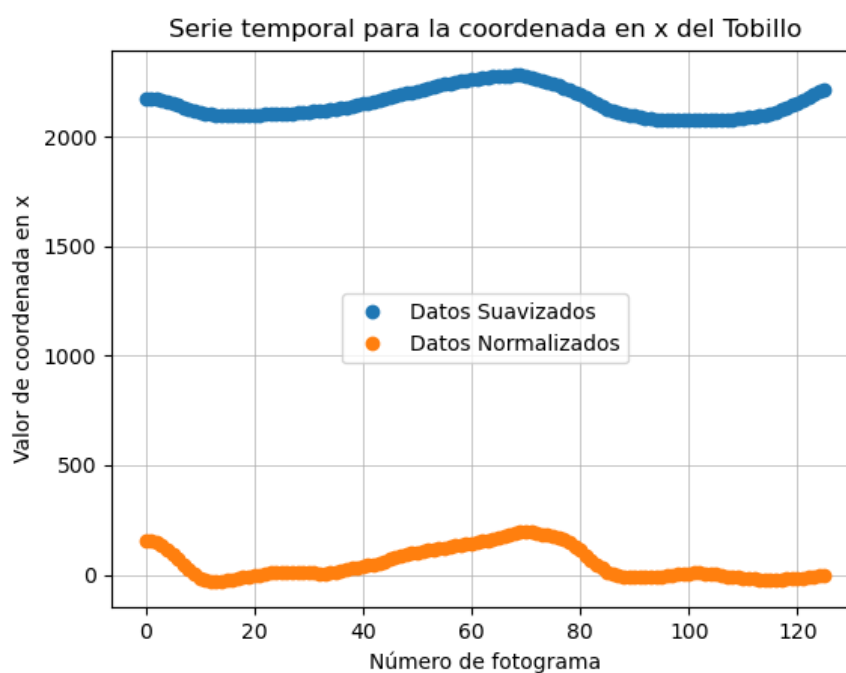


Figura 4.15: Serie temporal correspondiente a la coordenada en  $x$  del tobillo derecho, Vista frontal.

Este efecto se puede apreciar de manera más clara al comparar la Figura 4.14(a) con la Figura 4.14(b). En la Figura 4.14(a), se muestran los esqueletos extraídos en dos fotogramas distintos antes de aplicar la normalización, lo cual revela los cambios que se producen en la vista frontal a medida que el deportista se acerca a la cámara. Por otro lado, en la Figura 4.14(b), la normalización caracteriza los esqueletos de manera independiente a su posición y los presenta en una escala común. Este proceso también se aplica en el plano sagital y resulta de gran ayuda para el cálculo del SGEI.

### 4.3. Extracción de las características biomecánicas

La extracción de características biomecánicas implica identificar y cuantificar los patrones de movimiento de los deportistas, teniendo en cuenta variables como ángulos de articulaciones, velocidades, distancias, entre otros. Estos datos reflejan la dinámica de la técnica de los patinadores y se utilizan como entrada a los sistemas de clasificación para determinar el tipo de empuje utilizado por el patinador. Es importante resaltar que para el análisis biomecánico se toma en cuenta el sistema BIOMIN-PATIN específicamente en su módulo dinámico visto en la Sección 2.2.

Se utilizan dos enfoques para la extracción de características biomecánicas debido a sus fortalezas y aplicaciones: el enfoque gráfico, donde se obtienen imágenes SGEI para cada patinador que se clasifican mediante una CNN, y el enfoque de datos, conformado por un vector de características como se puede visualizar en el Apéndice B. Donde se obtiene un conjunto de valores para cada individuo analizado, como ángulos máximos, ángulos mínimos, velocidades y distancias. A continuación, se procede a implementar un clasificador mediante Support Vector Machine (SVM). La combinación de estos enfoques permite obtener una comprensión más completa y precisa de las características biomecánicas de los deportistas.

#### 4.3.1. Obtención del SGEI

Se generan imágenes SGEI a partir de secuencias de siluetas binarias de esqueletos obtenidas de cada fotograma del video capturado. Las siluetas binarias son obtenidas al graficar y unir los puntos correspondientes a las partes clave del cuerpo en un determinado fotograma. En la Figura 4.16 se muestra un ejemplo de una secuencia frontal, donde se aprecia el movimiento del individuo a través de los esqueletos reconstruidos. Además, la Figura 4.17 presenta un ejemplo de la vista sagital. Ambas figuras siguen un orden de movimiento de izquierda a derecha y de arriba hacia abajo.

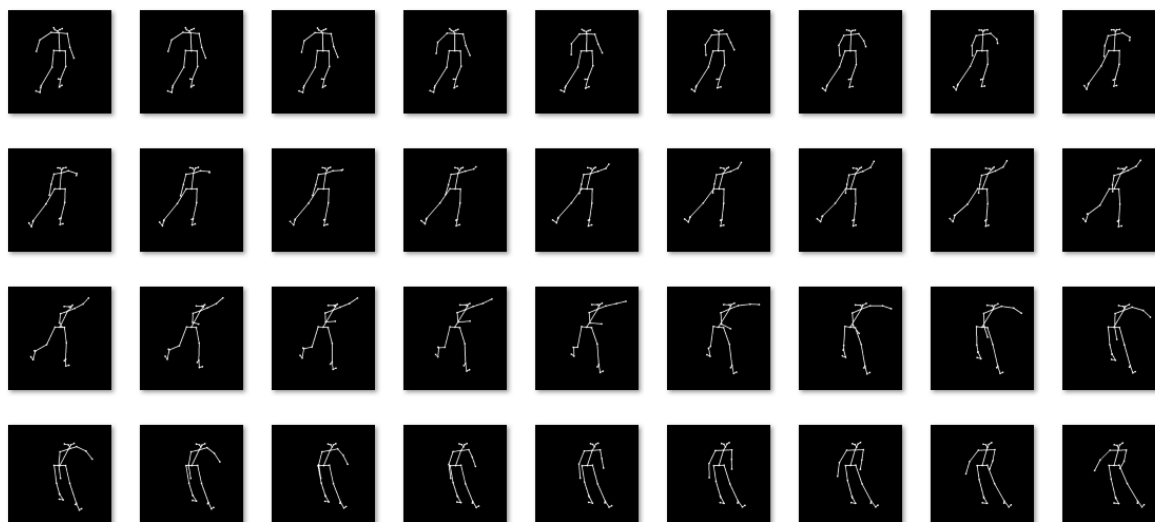


Figura 4.16: Secuencia de siluetas binarias frontales.

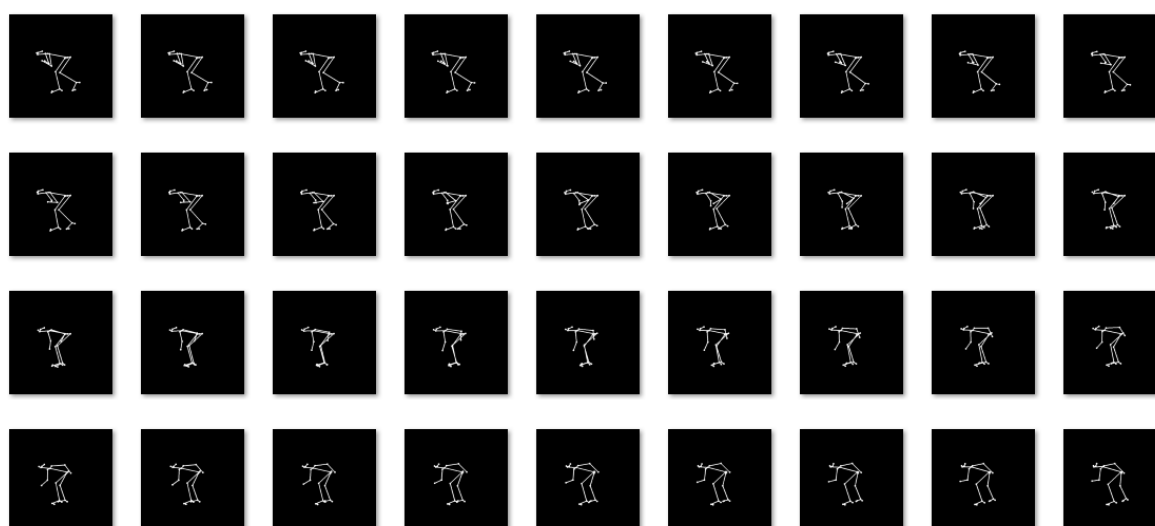


Figura 4.17: Secuencia de siluetas binarias sagitales.

Utilizando estas secuencias de esqueletos, se calcula el SGEI para cada archivo de video y persona correspondiente, siguiendo (2.5). Dado que cada imagen correspondiente a una silueta binaria tiene un tamaño de 600 x 600 píxeles, el SGEI obtenido también tendrá las mismas dimensiones que posteriormente se transformaran a 224 x 224 píxeles para la entrada del clasificador correspondiente. Los resultados se ilustran en las Figuras 4.18(a) y 4.18(b). El SGEI combina y resume la información y el movimiento del deportista en una sola imagen, lo que permite visualizar las características



distintivas de cada tipo de empuje.

En las Figuras 4.18(a) y 4.18(b), las zonas más claras indican menor movimiento, es decir, indican menor actividad o cambios en esa área, mientras que las zonas más difusas corresponden a partes del cuerpo que experimentan un mayor movimiento, como los brazos y las piernas, en comparación con las partes centrales del tronco, que se muestran más iluminadas. Estas observaciones revelan características generales de la técnica de patinaje, como la postura central del individuo, que se destaca cuando las piernas y los brazos están extendidos durante el desplazamiento.

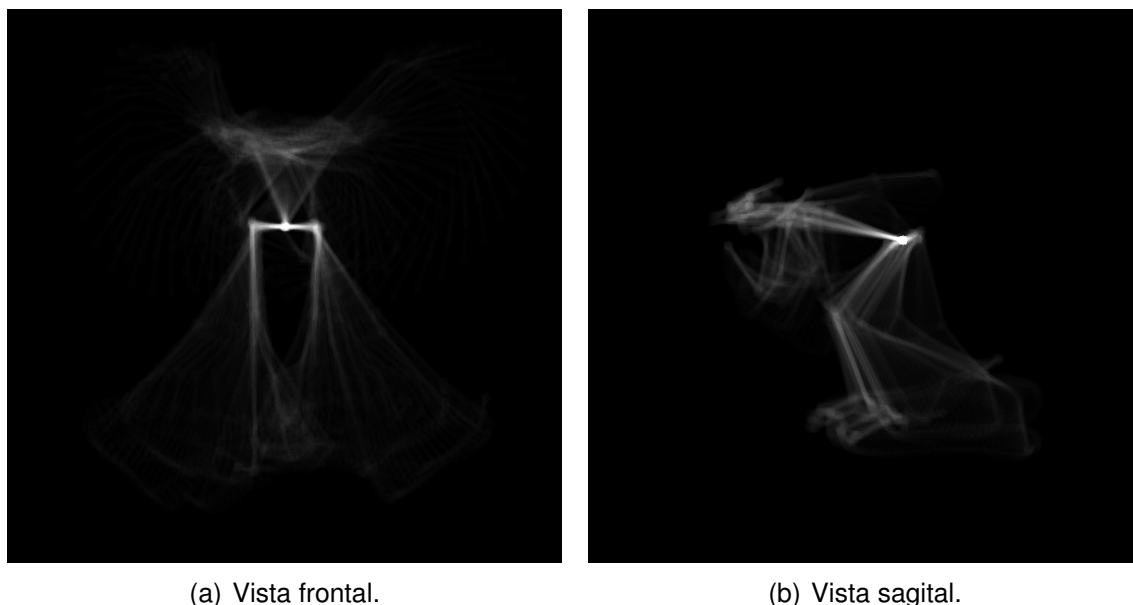


Figura 4.18: Imágenes SGEI generadas a partir de siluetas binarias

### 4.3.2. Cálculo de variables espacio-temporales

Durante la ejecución de un empuje en línea recta, se producen cambios e interacciones entre variables cinemáticas relevantes para los deportistas. Estas variables brindan información valiosa sobre las características del movimiento y son medidas importantes para clasificar el tipo de técnica utilizada. Al calcular estas variables, es necesario considerar aspectos específicos del patinaje en línea recta. Independientemente del tipo de empuje utilizado, esta técnica se compone de dos periodos: uno derecho y otro izquierdo. En el cálculo, se trata la ejecución total de ambos periodos como un ciclo, por lo tanto, en el cálculo de las variables cinemáticas, un medio ciclo se refiere a un periodo derecho o izquierdo. A continuación, se presentan las variables

espacio-temporales cuantificadas en este estudio.

- **Distancia recorrida durante medio ciclo:** Distancia promedio recorrida por el patinador durante la ejecución de los periodos derecho e izquierdo.
- **Distancia recorrida durante un ciclo entero:** Distancia promedio recorrida durante la ejecución de un ciclo completo.
- **Tiempo de medio ciclo:** Promedio del tiempo de ejecución de un ciclo derecho o izquierdo.
- **Tiempo de ciclo entero:** Promedio del tiempo que el patinador demora en ejecutar un ciclo total.
- **Velocidad:** Velocidad promedio de desplazamiento del patinador, es decir, la relación entre la distancia recorrida y el tiempo transcurrido.
- **Amplitud lateral máxima de piernas:** Medida de la máxima separación horizontal entre los tobillos del patinador en vista sagital.
- **Distancia máxima y mínima de hombro a rodilla:** Medida de la distancia euclidiana máxima entre el punto del hombro y la rodilla, tanto en el lado izquierdo como en el lado derecho del cuerpo.
- **Amplitud frontal máxima y mínima de piernas:** Medida de la máxima separación horizontal entre los tobillos del patinador en vista frontal.
- **Distancia entre hombros:** Distancia entre los hombros izquierdo y derecho.
- **Separación máxima del suelo de pie derecho:** Medida de la máxima separación alcanzada por el pie derecho respecto al suelo durante la fase de recuperación.
- **Separación máxima del suelo de pie izquierdo:** Medida de la máxima separación alcanzada por el pie izquierdo respecto al suelo durante la fase de recuperación.

En la Figura 4.19 se puede apreciar la representación de algunas de estas mediciones.

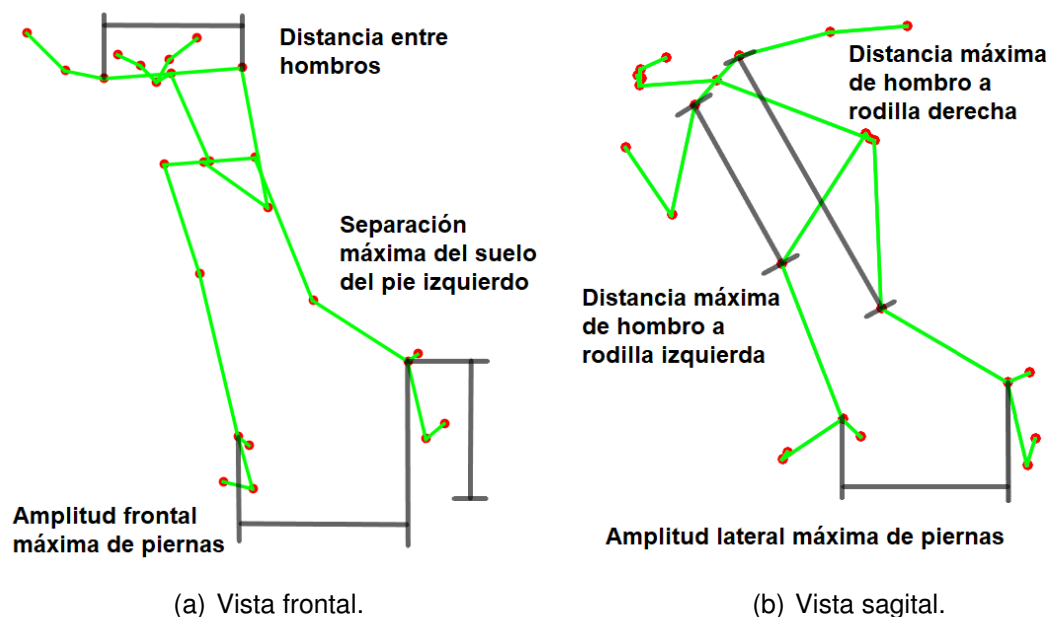


Figura 4.19: Indicación de las ubicaciones en las que se miden las variables espacio-temporales.

### 4.3.3. Cálculo de ángulos

Dada la estructura anatómica y la composición corporal de los seres humanos, es posible observar que varias uniones de segmentos o extremidades articuladas forman ángulos entre sí. Estos ángulos varían cuando se realiza cualquier tipo de movimiento corporal, ya que las extremidades y articulaciones involucradas también modifican su posición. En consecuencia, estos ángulos proporcionan información significativa sobre la dinámica del movimiento realizado.

Con las coordenadas de los puntos clave, una vez que han sido sometidas al pre-procesamiento de datos explicado en la Sección 4.2, se procede a calcular los ángulos principales que varían durante la ejecución de la técnica de patinaje en línea recta. Estos ángulos se obtienen tanto en la vista frontal como en la vista sagital. La Tabla 4.1 presenta los ángulos correspondientes a la vista frontal, mientras que la Tabla 4.2 muestra los ángulos de la vista lateral. Además, las Figuras 4.20(a), 4.20(b), 4.21(a) y 4.21(b) ilustran los ángulos a calcular y sus respectivos identificadores. En las Tablas 4.1 y 4.2, los identificadores de los puntos de *OpenPose* que están acompañados de la letra *h* indican que se ha realizado una proyección horizontal de dicho punto, mientras

que la letra  $v$  indica una proyección vertical. Estas proyecciones son necesarias para encontrar el ángulo requerido y se definen en (4.3) y (4.4).

$$P_{hi} = (x_i, y_i) \pm (100, 0) \quad (i = 0, 1, 2, \dots, 24) \quad (4.3)$$

$$P_{vi} = (x_i, y_i) \pm (0, 100) \quad (i = 0, 1, 2, \dots, 24) \quad (4.4)$$

Tabla 4.1: Ángulos para la vista frontal

Identificador	Detalle	Puntos clave de <i>OpenPose</i>
a	Pie externo derecho	(11h,11,22)
b	Pie externo izquierdo	(14h,14,19)
c	Pie interno derecho	(11h,11,22)
d	Pie interno izquierdo	(14h,14,19)
e	Empuje hombro interno derecho	(2v,2,22)
f	Empuje hombro interno izquierdo	(5v,5,19)
g	Empuje tobillo derecho	(10,11,11h)
h	Empuje tobillo izquierdo	(13,14,14h)
i	Empuje hombro externo derecho	(2v,2,22)
j	Empuje hombro externo izquierdo	(5v,5,19)
k	Entrepiera	(10,8,13)

Tabla 4.2: Ángulos para la vista sagital

Identificador	Detalle	Puntos clave de <i>OpenPose</i>
a	Rodilla derecha	(9,10,11)
b	Rodilla izquierda	(12,13,14)
c	Canilla derecha	(10,11,11h)
d	Canilla izquierda	(13,14,14h)
e	Cadera derecha	(2,9,10)
f	Cadera izquierda	(5,12,13)
g	Hombro derecho	(3,2,9)
h	Hombro izquierdo	(7,5,12)
i	Tronco	(1,8,8h)
j	Tobillo derecho	(10,11,22)
k	Tobillo izquierdo	(13,14,19)
l	Entrepiera	(10,8,13)
m	Codo derecho	(2,3,4)
n	Codo izquierdo	(5,6,7)

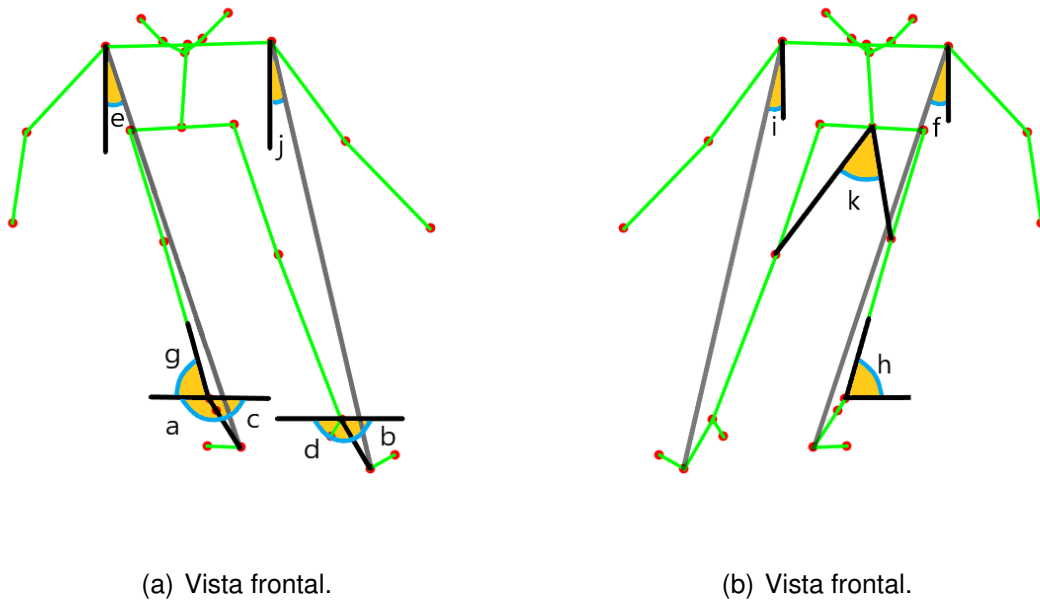


Figura 4.20: Ángulos calculados en vista frontal.

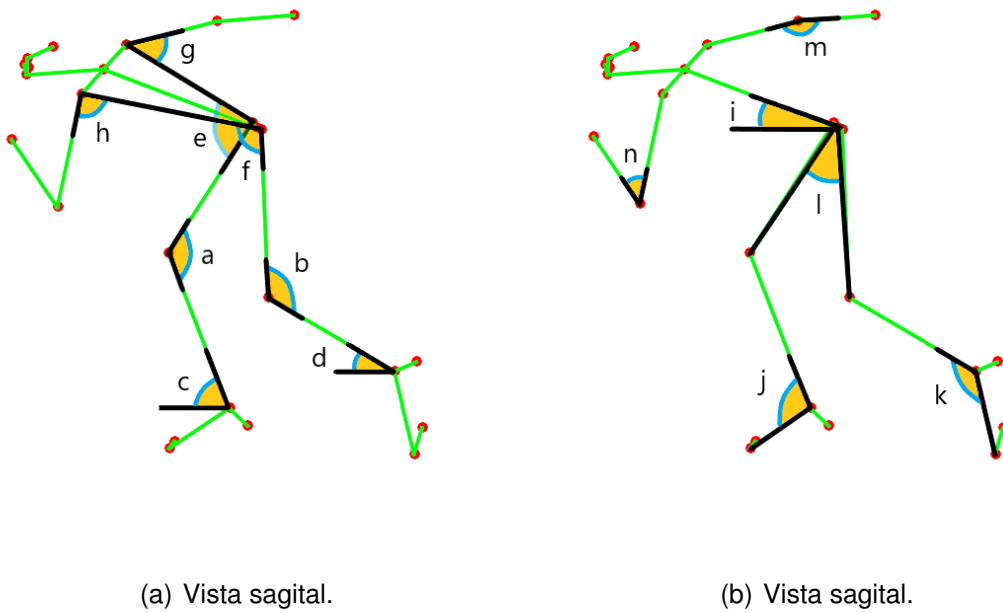


Figura 4.21: Ángulos calculados en vista sagital.

#### 4.4. Métodos de clasificación

A continuación, se presenta la implementación de los métodos de clasificación utilizados en este trabajo. Los códigos correspondientes a esta implementación, junto con

las diversas formas de obtención de las representaciones de imágenes y los vectores de características, están disponibles en el repositorio de código fuente [91].

#### 4.4.1. Clasificación con CNN

Una vez que se ha realizado el procesamiento de datos y se ha calculado el SGEI se procede a clasificarlos mediante una CNN presentada en la Sección 2.5.1. Este procedimiento se ilustra en la Figura 4.22. Para la extracción de características y patrones relevantes, además de su posterior clasificación, se escogió el modelo Visual Geometry Group (VGG19) presentado en la Sección 2.5.1.1. VGG19 es un modelo ampliamente utilizado por su precisión en la clasificación de imágenes, además de tener una arquitectura simple y fácil de entender. La implementación de este modelo se realiza usando la biblioteca Keras de TensorFlow en python.

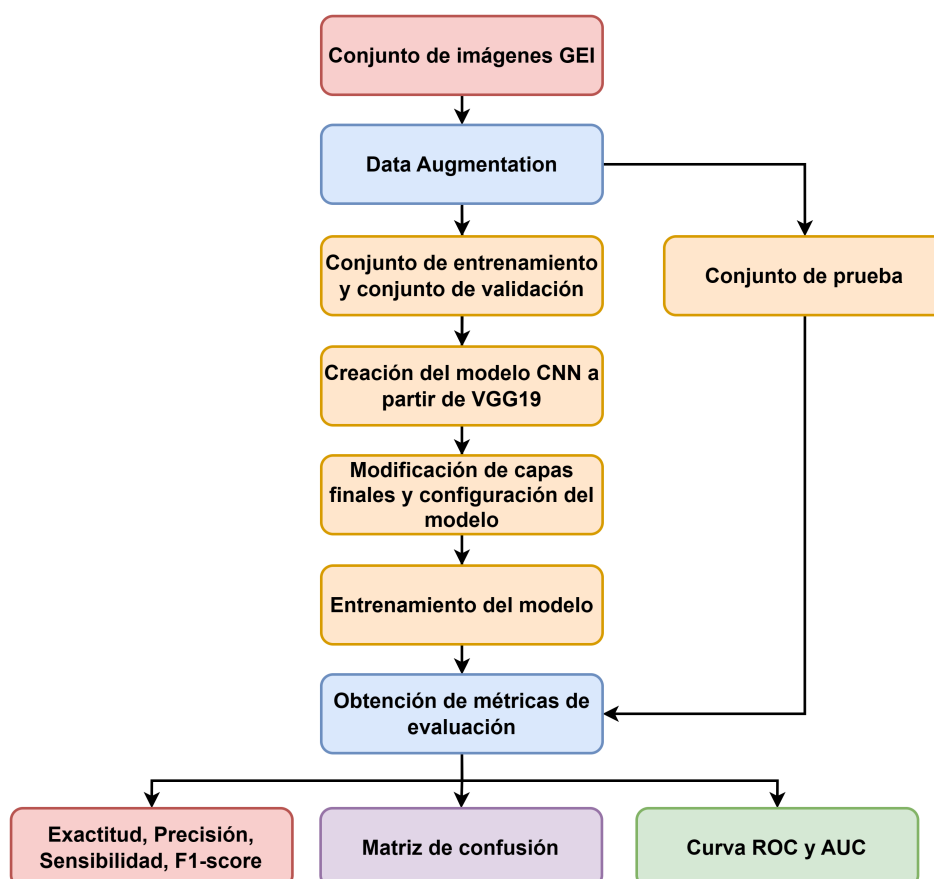


Figura 4.22: Diagrama de flujo para la clasificación con CNN.

Debido a que la base de datos obtenida se caracteriza por contener un número limitado de registros, se optó por el uso de *Data augmentation*, que es una técnica utili-

zada en el campo de ML. Consiste en aplicar transformaciones y manipulaciones a los datos existentes, tal como rotación, desplazamiento, ampliación o reducción a las imágenes ingresadas. De esta manera, se logra generar información nueva de manera realista y representativa. Al aumentar datos de manera aleatoria, se busca reducir el riesgo de sobreajuste y mejorar la capacidad de generalización del modelo. Además, se utiliza la transferencia de aprendizaje aprovechando un modelo previamente entrenado con la base de datos *Imagenet* [92]. Esta técnica consiste en aprovechar el conocimiento adquirido y la capacidad de generalizar el modelo pre-entrenado. Mediante el ajuste de los pesos, se logra adaptar el conocimiento adquirido a la tarea específica que se desea abordar. Como resultado, la transferencia de aprendizaje con la base de datos *Imagenet* tiene un mejor rendimiento, además de acelerar el proceso de entrenamiento.

Durante el entrenamiento, las primeras capas mantienen sus pesos, debido a que detectan los patrones comunes, mientras que las capas posteriores ajustan sus pesos con la información previa, estos se adaptan con mayor facilidad, permitiéndoles detectar características más específicas de la base de datos de entrada.

El modelo de la CNN consta de varios bloques de filtros pequeños, con capas regularizadas para evitar el sobreajuste. La última capa es completamente conectada y se establece el número de neuronas (canales permitidos por capa). También se aplica una técnica de regularización llamada *Dropout*. Consiste en desactivar aleatoriamente un grupo de neuronas en una capa determinada durante cada paso de época del entrenamiento. Al desactivar estas neuronas, se evita que las neuronas individuales dependan excesivamente de otras neuronas específicas para realizar predicciones. Esta técnica ayuda a evitar el sobreajuste al impedir que el modelo dependa demasiado de un subconjunto particular de características o conexiones.

#### 4.4.2. Clasificación con SVM

Se aplica el método de SVM descrito en la Sección 2.5.2 para clasificar un conjunto de vectores de características. Este procedimiento, en conjunto con la aplicación de Random Forest (RF), se ilustra en la Figura 4.23.

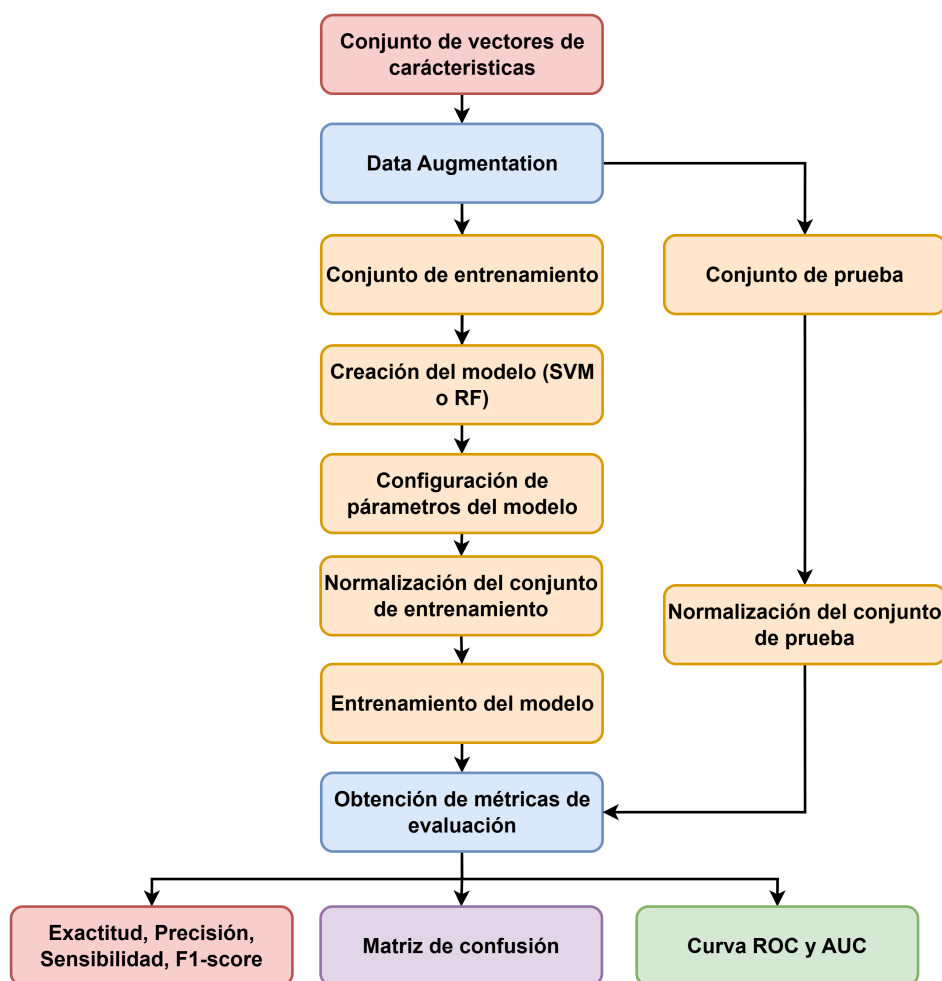


Figura 4.23: Diagrama de flujo para la clasificación con SVM y RF.

Estos vectores se generan con el resultado de los cálculos de variables espacio-temporales descritas en la Sección 4.3.2, que abarcan distancias, tiempos y velocidades, así como los ángulos del cuerpo obtenidos a través de principios de trigonometría y relaciones espaciales de la Sección 4.3.3. En total, se generan 95 variables que se utilizan como características de los vectores indicados en el Apéndice B. Las características relacionadas con los ángulos incluyen valores máximos, mínimos y su valor promedio. Algunas de las variables espacio-temporales también especifican valores máximos y mínimos.

A continuación, el conjunto de vectores se utiliza como entrada para el algoritmo SVM. La implementación de este modelo se realiza en Python utilizando la biblioteca *scikit-learn* (*sklearn*). Además, se emplean los parámetros de regularización  $C$  y  $\gamma$  para evitar el sobreajuste durante el entrenamiento, en conjunción con la elección de



un kernel Radial Basis Function (RBF).

#### 4.4.3. Clasificación con RF

Para implementar el algoritmo de RF, se utiliza el mismo conjunto de datos y la misma división de entrenamiento y prueba que en el sistema SVM, ya que ambos algoritmos clasifican información en forma de vectores de características. En el caso de RF, se emplean los parámetros `n_estimators` con un valor predeterminado de 100 para indicar el número de árboles en el algoritmo, pues así se consigue un equilibrio razonable entre rendimiento y precisión. Se utiliza `random_state` un valor de 1 para asegurar la replicabilidad de los resultados obtenidos.

## Capítulo 5 Resultados

En este capítulo se presentan los resultados al emplear diferentes métodos de clasificación. Se realiza un análisis, procesamiento y evaluación de la eficiencia de clasificación de la información recopilada de los 19 deportistas voluntarios que participaron en el estudio.

### 5.1. Análisis

En esta sección se presentan los resultados de la biomecánica de los patinadores. En la Sección 5.1.1, se proporcionan valores de los parámetros espacio-temporales generados, así como los ángulos obtenidos en vista frontal y sagital. Por otro lado, en la Sección 5.1.2 se analizan los patrones de los ángulos del tren inferior de cada tipo de empuje realizado por los patinadores. La evaluación y precisión de los clasificadores se abordan en la Sección 5.1.3. Finalmente, en la Sección 5.1.3.4 se presenta una comparación entre los clasificadores utilizados.

#### 5.1.1. Parámetros espacio-temporales y ángulos

En las Tablas 5.1, 5.2 y 5.3, se muestra el promedio y la desviación estándar ( $\pm$ ), de los parámetros espacio-temporales y ángulos, en relación al total de participantes del estudio. En estas tablas, la etiqueta (D) corresponde al lado derecho del cuerpo y (I) al lado izquierdo. Las Tablas 5.2 y 5.3 muestran los promedios correspondientes a los valores medios de los ángulos. Las tablas completas con los valores medios, máximos y mínimos de los ángulos se presentan en el Apéndice B.

Estos datos se han calculado con base en el conjunto completo de vectores de características, los cuales a su vez se derivaron de las coordenadas de los puntos clave extraídos de los videos capturados de los 19 deportistas voluntarios mientras llevaban a cabo uno de los 3 tipos de empuje.

Los resultados muestran una notable similitud con investigaciones previas de patinaje llevadas a cabo por [15, 25], lo cual valida los datos obtenidos mediante el sistema desarrollado. Por ejemplo, [25] indica ángulos de rodilla mínimos en vista sagital entre el

rango de  $95.7^\circ$  a  $98.6^\circ$ , mientras que [15] indica velocidades en el rango de 8.225 m/s y 11.2 m/s. Estos datos tienen concordancia con los datos en la Tabla 5.1 y los datos del Apéndice B. En estos trabajos los ángulos son medidos manualmente, pausando el vídeo y graficando líneas sobre el deportista y midiendo el ángulo obtenido.

Se observa que la velocidad del doble empuje (9.34 m/s) es superior a la velocidad de empuje clásico (8.26 m/s) y pendular (8.62 m/s), a pesar de que se recorre distancias similares, (clásico: 13.43 m, doble: 14.57 m y pendular: 14.68 m). Sin embargo, el doble empuje recorre estas distancias en menor tiempo (clásico: 1.66 s, doble: 1.6 s, y pendular: 1.75 s). Al ajustar las distancias a una misma escala para facilidad de comparación, se obtienen valores de tiempo y velocidad en los distintos empujes: clásico (1.82 s y 8.06 m/s), doble (1.61 s y 9.12 m/s) y pendular (1.75 s y 8.38 m/s)). Esto confirma que doble empuje continúa teniendo el menor tiempo entre los diferentes tipos de empujes.

Por consiguiente, se demuestra que el doble empuje tiene mayor eficiencia en comparación con los otros tipos de empuje, debido a que este empuje logra una mayor velocidad con menor esfuerzo, pues se da una utilización simultánea de ambos pies y una mejor sincronización de los movimientos para lograr una mayor propulsión y velocidad, lo que indica una utilización más eficiente de la energía y mayor eficacia de rendimiento. Además, se observa que la amplitud mínima frontal del empuje clásico (D de 0,14 m e l de 0,11 m), doble (D de 0.11 m e l de 0.13 m) y pendular (D de 0.17 m e l de 0.17 m) es menor que la distancia de los hombros, según se establece en la Sección 2.1.2 correspondiente a la fase 1 denominada posición básica.

Tabla 5.1: Distancias, tiempos y velocidades (promedio  $\pm$  desviación estándar)

<b>Características</b>	<b>Clásico</b>	<b>Doble</b>	<b>Pendular</b>
Max Amplitud lateral D (m)	0,5 $\pm$ 0,12	0,55 $\pm$ 0,12	0,55 $\pm$ 0,13
Max Amplitud lateral I (m)	0,43 $\pm$ 0,08	0,45 $\pm$ 0,11	0,44 $\pm$ 0,1
Tiempo medio ciclo D (s)	0,75 $\pm$ 0,23	0,71 $\pm$ 0,22	0,81 $\pm$ 0,29
Tiempo medio ciclo I (s)	0,82 $\pm$ 0,3	0,83 $\pm$ 0,22	0,9 $\pm$ 0,31
Tiempo ciclo completo (s)	1.66 $\pm$ 0.62	1.6 $\pm$ 0.41	1.75 $\pm$ 0.51
Distancia medio ciclo D (m)	6,08 $\pm$ 1,56	6,5 $\pm$ 1,65	6,76 $\pm$ 1,97
Distancia medio ciclo I (m)	6,64 $\pm$ 2,17	7,59 $\pm$ 1,62	7,51 $\pm$ 2,13
Distancia ciclo (m)	13.43 $\pm$ 4.34	14.57 $\pm$ 2.55	14.68 $\pm$ 3.27
Velocidad (m/s)	8.26 $\pm$ 0.83	9.34 $\pm$ 1.26	8.62 $\pm$ 1.09
Max distancia hombro rodilla D (m)	0.63 $\pm$ 0.08	0.63 $\pm$ 0.08	0.69 $\pm$ 0.07
Max distancia hombro rodilla I (m)	0.67 $\pm$ 0.06	0.72 $\pm$ 0.06	0.72 $\pm$ 0.06
Min distancia hombro rodilla D (m)	0.34 $\pm$ 0.05	0.36 $\pm$ 0.04	0.39 $\pm$ 0.04
Min distancia hombro rodilla I (m)	0.32 $\pm$ 0.04	0.37 $\pm$ 0.05	0.37 $\pm$ 0.05
Max amplitud frontal D (m)	0,44 $\pm$ 0,09	0,5 $\pm$ 0,09	0,49 $\pm$ 0,09
Max amplitud frontal I (m)	0,39 $\pm$ 0,07	0,48 $\pm$ 0,1	0,41 $\pm$ 0,07
Min amplitud frontal D (m)	0,14 $\pm$ 0,06	0,11 $\pm$ 0,05	0,17 $\pm$ 0,05
Min amplitud frontal I (m)	0,11 $\pm$ 0,08	0,13 $\pm$ 0,05	0,17 $\pm$ 0,07
Media distancia hombros (m)	0.24 $\pm$ 0.02	0.27 $\pm$ 0.02	0.26 $\pm$ 0.02
Max separación pie suelo D (m)	0.19 $\pm$ 0.06	0.19 $\pm$ 0.05	0.25 $\pm$ 0.08
Max separación pie suelo I (m)	0.2 $\pm$ 0.06	0.18 $\pm$ 0.05	0.26 $\pm$ 0.07

Los ángulos promedio del exterior del pie D (89.04°) e I (82.82°) revelan una mayor inclinación en doble empuje, mientras que en el empuje clásico muestra una inclinación menor en comparación con los otros tipos de empuje. Como se observó en la Sección 2.1.3.2, el empuje pendular se caracteriza por realizar una tracción y empuje de manera no simultánea, así lo demuestra el parámetro de empuje interno (D de 12.92° e I de 9.57°) y externo (D de 20.31° e I de 19.49°). Sus respectivas desviaciones estándares son mayores que en los otros empujes debido al mayor rango de movimiento en la inclinación. Es importante tener en cuenta que los resultados pueden verse influenciados por la altura, peso, flexibilidad, fuerza y coordinación de los participantes. Por lo que cada individuo es único y puede responder de manera diferente a diferentes características. Por ejemplo, lo que puede ser una influencia positiva para un patinador puede no serlo para otro, y viceversa.

Tabla 5.2: Valores medios de los ángulos en vista frontal (promedio  $\pm$  desviación estándar)

<b>Características</b>	<b>Clásico (°)</b>	<b>Doble (°)</b>	<b>Pendular (°)</b>
Exterior pie D	84.71 $\pm$ 6.53	89.04 $\pm$ 4.7	88.91 $\pm$ 4.91
Exterior pie I	79.56 $\pm$ 4.53	82.82 $\pm$ 4.63	79.81 $\pm$ 5.25
Interior pie D	95.29 $\pm$ 6.53	90.96 $\pm$ 4.7	91.09 $\pm$ 4.91
Interior pie I	100.44 $\pm$ 4.53	97.18 $\pm$ 4.63	100.19 $\pm$ 5.25
Empuje interno D	10.71 $\pm$ 1.86	12.02 $\pm$ 2.03	12.92 $\pm$ 2.17
Empuje interno I	7.8 $\pm$ 1.86	8.99 $\pm$ 1.92	9.57 $\pm$ 2.4
Empuje externo D	18.72 $\pm$ 2.47	19.2 $\pm$ 1.74	20.31 $\pm$ 3.17
Empuje externo I	17.03 $\pm$ 2.5	19.71 $\pm$ 1.62	19.49 $\pm$ 2.69
Tobillo D	80.36 $\pm$ 4.27	82.04 $\pm$ 3.32	78.88 $\pm$ 4.79
Tobillo I	81.51 $\pm$ 4.22	84.15 $\pm$ 3.11	81.74 $\pm$ 4.4
Entrepiera	47.57 $\pm$ 6.24	46.58 $\pm$ 5.88	48.49 $\pm$ 5.85

Como se establece en la Sección 2.1.3.3, en el doble empuje se realiza una tracción y empuje de manera simultánea, gracias a estas características, no es necesario flexionar tanto las rodillas como se refleja en los valores promedio de rodilla D (118.66 °) e I (120.3 °). Además, estos valores presentan desviaciones más bajas que los otros tipos de empuje. Sin embargo, en este empuje la flexión del tronco (17.87 °) es mayor que en empuje clásico y pendular.

Tabla 5.3: Valores medios de los ángulos en vista sagital (promedio  $\pm$  desviación estándar)

<b>Características</b>	<b>Clásico (°)</b>	<b>Doble (°)</b>	<b>Pendular (°)</b>
Tronco	21.25 $\pm$ 4.19	17.87 $\pm$ 2.74	23.45 $\pm$ 4.64
Cadera D	86.63 $\pm$ 4.52	80.13 $\pm$ 3.93	87.42 $\pm$ 6.3
Cadera I	83.65 $\pm$ 6.24	81.15 $\pm$ 5.45	85.09 $\pm$ 8.73
Rodilla D	119.01 $\pm$ 7.1	118.66 $\pm$ 4.42	116.19 $\pm$ 7.31
Rodilla I	118.18 $\pm$ 5.71	120.3 $\pm$ 5.72	116.33 $\pm$ 6.53
Tobillo D	102.23 $\pm$ 8.62	101.2 $\pm$ 4.05	99.88 $\pm$ 5.84
Tobillo I	97.58 $\pm$ 5.06	101.1 $\pm$ 4.59	98.35 $\pm$ 6.41
Canilla D	53.8 $\pm$ 6.9	57.74 $\pm$ 3.13	52.25 $\pm$ 4.99
Canilla I	54.85 $\pm$ 5.9	55.65 $\pm$ 6.09	54.14 $\pm$ 6.78
Entrepiera	31.29 $\pm$ 7.46	31.18 $\pm$ 6.91	33.1 $\pm$ 7.34
Codo D	125.94 $\pm$ 18.84	132.01 $\pm$ 17.05	116.27 $\pm$ 25.44
Codo I	141.96 $\pm$ 14.82	129.6 $\pm$ 21.5	141.2 $\pm$ 14.75
Hombro D	51.57 $\pm$ 12.15	41.24 $\pm$ 15.23	53.07 $\pm$ 17.62
Hombro I	33.83 $\pm$ 10.14	27.8 $\pm$ 12.82	25.17 $\pm$ 17.43

Los parámetros espacio-temporales y ángulos presentados en las tablas anteriores

evidencian una concordancia bilateral entre el medio ciclo derecho e izquierdo, pues los valores para el lado derecho del cuerpo no difieren de manera significativa con los valores correspondientes al lado izquierdo. Esto sugiere un equilibrio biomecánico adecuado en la ejecución de los movimientos analizados. Además, se observa una sincronización y similitud significativas al analizar los parámetros medidos. Es decir, que indica que los movimientos y características de los miembros inferiores y superiores están coordinados y se asemejan entre sí. Estos hallazgos apoyan la idea de una distribución equitativa de peso y movimiento entre los miembros inferiores y superiores en ambos lados del cuerpo, lo que indica una ejecución equilibrada del patinaje de velocidad.

### 5.1.2. Comparación en los ángulos de las extremidades inferiores

A continuación, se presentan gráficas del ángulo en función del ciclo del patinaje de velocidad, obtenidas a partir de la flexión del tronco, rodilla, tobillo en vista sagital. La Figura 5.1 presenta los patrones generados, es decir, la forma de las curvas, de los tres tipos de empuje en la flexión realizada por el tronco.

La Figura 5.1 representa la flexión realizada por el tronco, se presenta los patrones generados por los tres tipos de empuje. Las regiones de colores representan las variaciones de los ángulos producidos por los 19 deportistas voluntarios, mientras que las líneas de color oscuro representan el promedio. Las Figuras 5.2 y 5.3 siguen el mismo esquema, donde las gráficas de color rojo corresponden al empuje clásico, para el doble empuje son las gráficas de color verde y por último las gráficas de color azul representan al empuje pendular.

Se observa que los patrones entre los diferentes tipos de empujes son muy similares, sin embargo, existen diferencias en el rango de flexión en distintas partes del ciclo. En la Figura 5.1 se puede apreciar que en la fase de recuperación, que corresponde del 40% al 80% del ciclo, existe una mayor flexión del tronco. Esto se debe al traslado del peso que ocurre desde la fase de empuje. Mientras que en el empuje clásico la fase de recuperación se traslada al 42% del ciclo, en el empuje pendular se presentan valores menos definidos, con un rango mayor de movimiento del tronco.

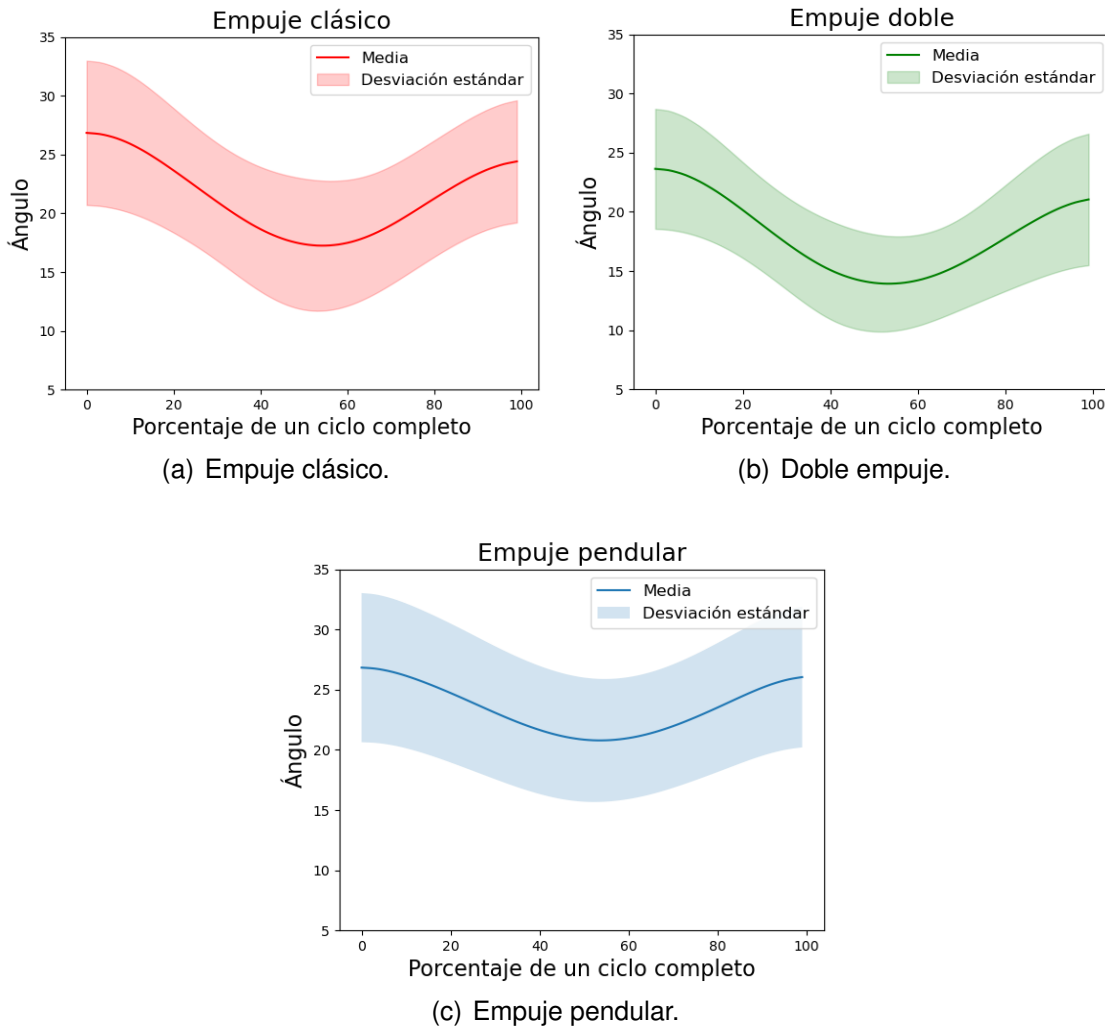


Figura 5.1: Flexión del tronco

En la Figura 5.2 se distingue los patrones en los ángulos articulares de la rodilla izquierda. Se observa una flexión, es decir, una disminución en el valor del ángulo, en la fase de empuje que corresponde al 20% de ciclo, donde la mayor parte del peso del cuerpo está apoyado en la pierna izquierda, lo que provoca una flexión máxima en la fase de recuperación, que abarca del 25% al 80% del ciclo. El último, el 20% corresponde a la fase final, donde la pierna vuelve a su estado inicial de apoyo. Se observa que en el empuje clásico y pendular se realiza una mayor flexión de rodilla que en el doble empuje.

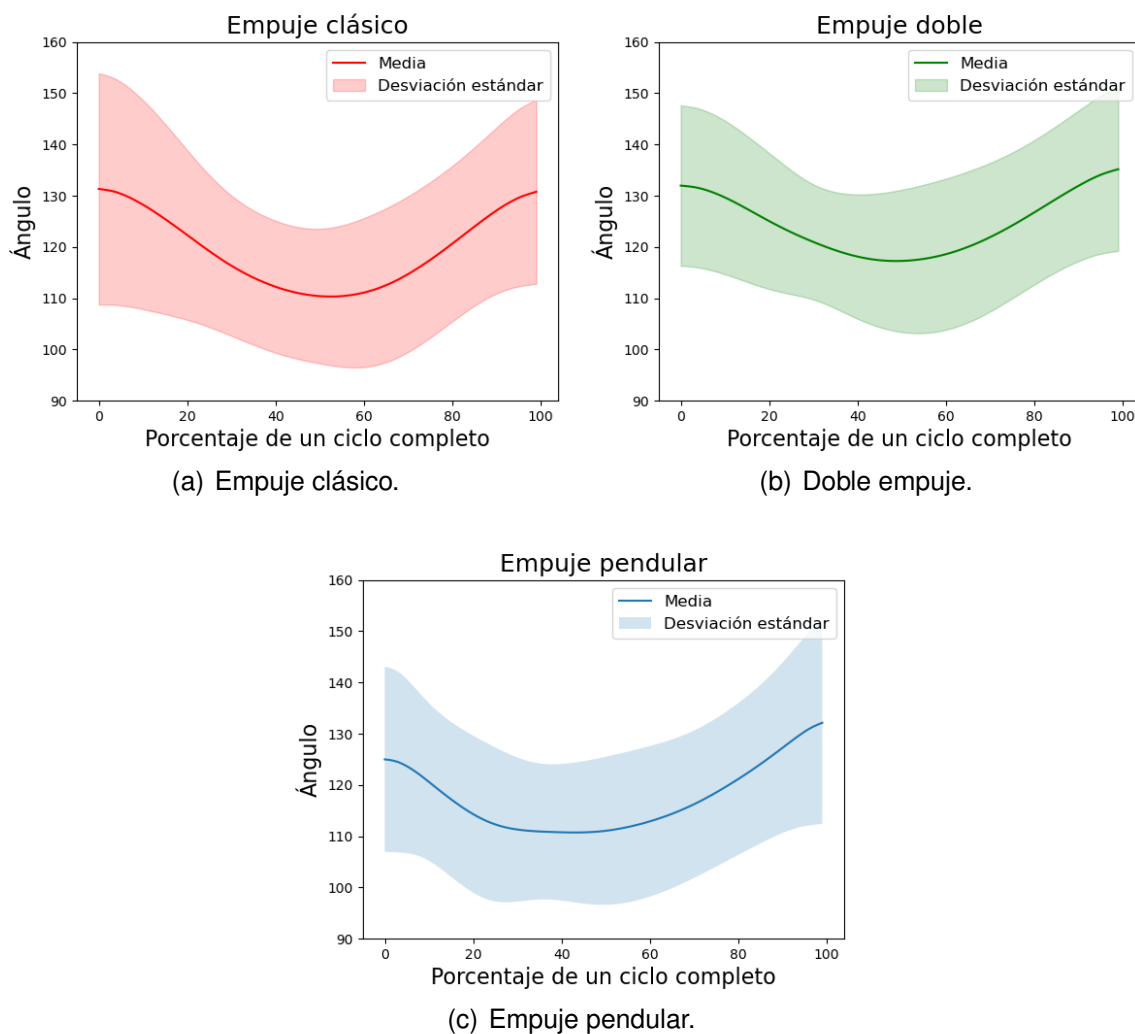


Figura 5.2: Flexión de rodilla izquierda.

En la Figura 5.3 se presenta el patrón generado por el ángulo del tobillo derecho. Se observa un patrón más definido en doble empuje, esto es, que el valor de la desviación estándar es más constante, debido a su movimiento constante de inclinación. Por otro lado, el empuje clásico muestra una flexión aplanada, alcanzando máximo de 125 grados. En contraste, el empuje pendular muestra un patrón que combina características del empuje clásico y el doble empuje.



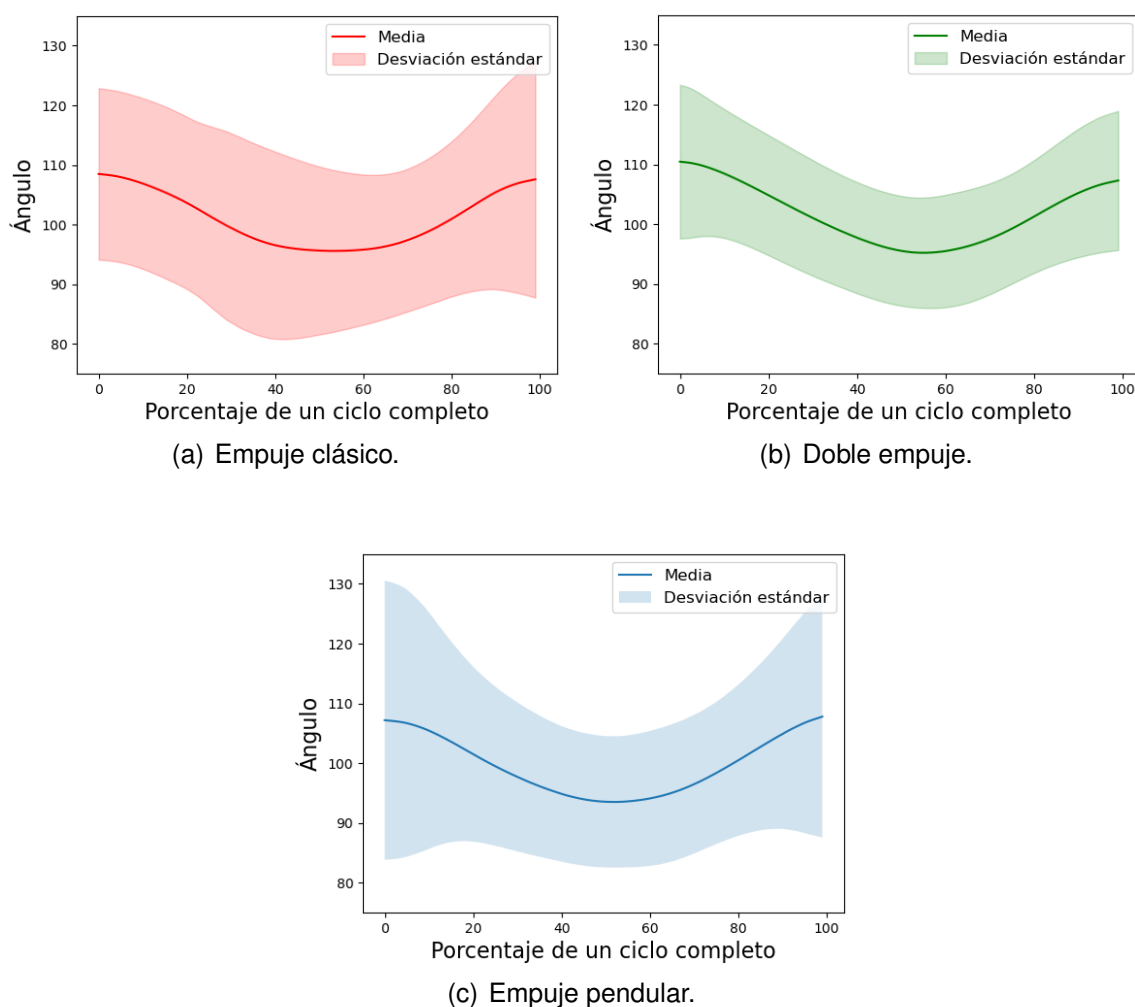


Figura 5.3: Flexión de tobillo derecho.

### 5.1.3. Evaluación de los clasificadores

A continuación, se presenta los resultados de clasificación de Convolutional Neural Network (CNN) y Support Vector Machine (SVM) respectivamente de los 3 tipos de empuje de un patinador identificados numéricamente como clásico (0), doble(1), pendular(2), utilizando herramientas para evaluar la precisión de los clasificadores como son la matriz de confusión, la curva Receiver Operating Characteristic (ROC), exactitud, precisión, sensibilidad y F1-score.

### 5.1.3.1. Clasificación usando CNN

Para la clasificación mediante imágenes, se utiliza CNN como se mencionó en la Sección 4.3.1, como entrada se usa imágenes Skeleton Gait Energy Image (SGEI) en el plano frontal. Un total de 2618 imágenes SGEI fueron utilizadas para la clasificación, la cual se llevó a cabo de forma manual por un experto. Estas imágenes se dividieron en 759 de empuje clásico, 963 de doble empuje y 896 de empuje pendular. El 70 % de cada tipo de empuje se utiliza en entrenamiento, 15 % en validación y 15 % en pruebas.

Para evitar que el modelo lea todas las imágenes a la vez, se escoge un *batch size* de 32, esto es, que las imágenes se procesan en lotes de 32 para evitar el consumo excesivo de memoria. El valor de 32 es un valor predeterminado en varias librerías que manejan procesos de clasificación con imágenes. Se emplea un modelo con el optimizador Stochastic Gradient Descent (SGD) utilizando una tasa de aprendizaje de 0.002 con el objetivo de acelerar el entrenamiento del modelo. Si bien la velocidad de entrenamiento puede ser relevante desde una perspectiva práctica o de eficiencia, no tiene un impacto directo en la precisión de clasificación del modelo. Además, se usa la función de pérdidas *categorical cross entropy*, que es utilizada comúnmente para clasificaciones multiclase, pues compara la distribución de probabilidad predicha por el modelo con la distribución real, y penaliza las predicciones incorrectas.

Como primeros resultados de la clasificación CNN se tiene los valores de exactitud en entrenamiento de 80.97 % y en el conjunto de validación una exactitud de 88.92 %. En la Figura 5.4 se presenta la evolución de las curvas de pérdidas y exactitud. La curva de pérdidas muestra el valor de la entropía cruzada calculada en la época correspondiente, dicho valor es mayor mientras mayor sea la discrepancia entre las predicciones y las clases reales. La curva de pérdidas muestra el valor de la entropía cruzada calculada en cada época del entrenamiento. Este valor decrece a medida que disminuye la discrepancia entre las predicciones del modelo y las clases reales. Se logra observar que la curva de pérdidas se aproxima a cero, mientras que la curva de exactitud alcanza un porcentaje cercano a 90 % que se considera alto, es decir que el número de épocas establecido es adecuado.

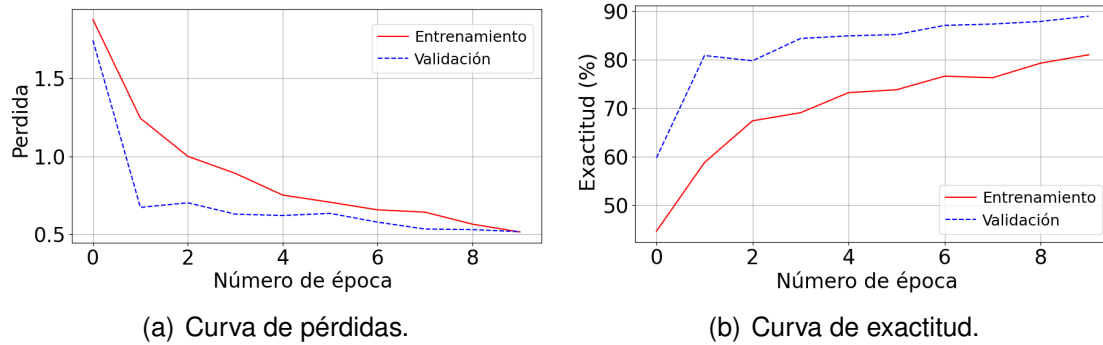


Figura 5.4: Curvas de entrenamiento.

Para la evaluación del conjunto de datos de prueba, se utilizó 377 imágenes SGEI. 91 de empuje clásico, 143 de doble empuje y 143 de pendular. En la Tabla 5.4 se presenta varias métricas de evaluación para cada tipo de empuje. A pesar de que se observa una disminución en la sensibilidad para la clase de empuje clásico, aún clasifica correctamente el 83.52 % de las imágenes de dicha clase, lo cual demuestra resultados prometedores.

Tabla 5.4: Evaluación de diferentes métricas para CNN.

<b>Tipos de empuje</b>	<b>Precisión (%)</b>	<b>Sensibilidad (%)</b>	<b>F1-score (%)</b>
Clásico	92.68	83.52	87.86
Doble	90.73	95.80	93.20
Pendular	89.58	90.21	89.90

Los resultados de la matriz de confusión y curvas ROC se presentan en las Figuras 5.5(a) y 5.5(b), presentando una exactitud de 90.72 % y un área bajo la curva de 98.24 % lo cual demuestra ser eficiente al momento de clasificar un conjunto de datos no visto anteriormente por el modelo.

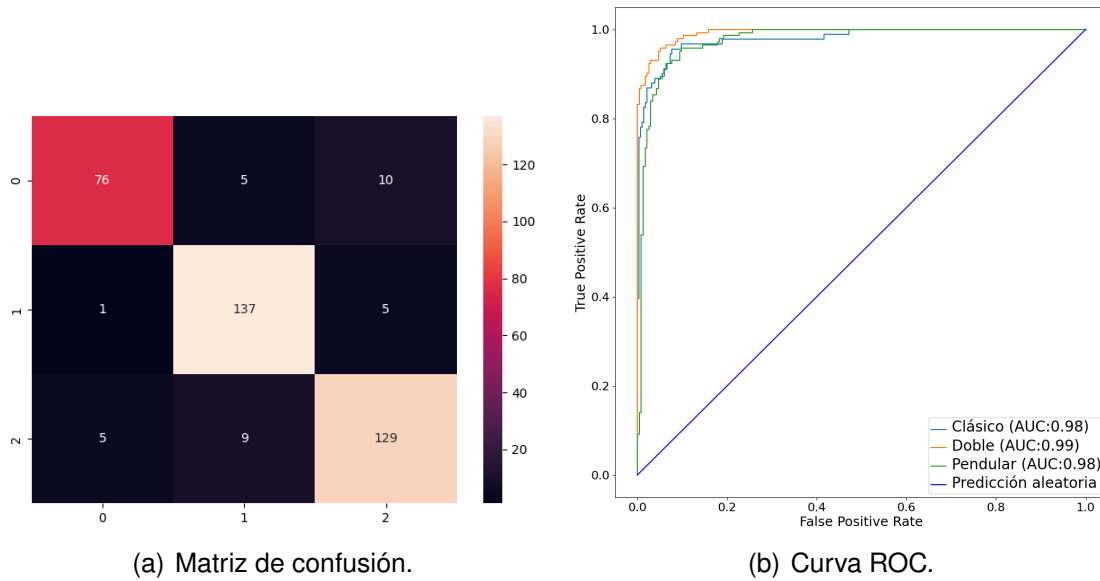


Figura 5.5: Evaluación de CNN.

### 5.1.3.2. Clasificación usando SVM

Según lo expuesto en la Sección 4.4.2, se procede a clasificar el conjunto de vectores de características utilizando el método de SVM con un kernel Radial Basis Function (RBF) debido a la naturaleza no lineal de los datos. El parámetro  $C$  se establece en 0.05, lo que resulta en un margen más amplio, y el parámetro Gamma se configura en 0.004, lo que incrementa la influencia de cada vector de entrenamiento en la construcción del hiperplano de decisión.

Se procesa un total de 264 vectores de características: 64 correspondientes a la clase de empuje clásico, 77 a la clase de doble empuje y 123 a la clase de empuje pendular. Estos datos se dividen en un conjunto de entrenamiento, que representa el 75 % del total, y un conjunto de pruebas, que constituye el 25 % restante. Para aumentar la cantidad de datos disponibles y equilibrar la distribución de muestras en cada clase, se aplica un proceso de generación de datos sintéticos. Como resultado, se obtienen 240 vectores de entrenamiento para cada clase. Los resultados del clasificador SVM son prometedores, mostrando un porcentaje de exactitud del 92.22 % en los datos de entrenamiento y del 93.94 % en el conjunto de vectores de prueba. El tiempo de ejecución para clasificación con SVM es de 0.692 segundos.

La Tabla 5.5 exhibe métricas adicionales de clasificación para el conjunto de datos de

prueba.

Tabla 5.5: Evaluación de diferentes métricas para SVM.

<b>Tipos de empuje</b>	<b>Precisión (%)</b>	<b>Sensibilidad (%)</b>	<b>F1-score (%)</b>
Clásico	84.21	100.00	91.43
Doble	95.00	100.00	97.44
Pendular	100.00	87.10	93.10

El modelo exhibe una capacidad precisa para clasificar las clases de empuje clásico y doble empuje, y aunque presenta una ligera disminución en la sensibilidad para la clase de empuje pendular, aún logra clasificar correctamente la gran mayoría de las muestras de dicha clase. Estos resultados positivos respaldan la viabilidad del modelo SVM con kernel RBF y los parámetros ajustados para la clasificación de este tipo de datos. La Figura 5.6 respalda estas observaciones al mostrar tanto la matriz de confusión como la curva ROC del algoritmo SVM. En la matriz de confusión se pueden apreciar las clases 0, 1 y 2 correspondientes a clásico, doble y pendular, respectivamente. Por otro lado, la curva ROC muestra que la clase correspondiente al empuje clásico alcanza un Area Under the ROC Curve (AUC) de 0.99, la clase de empuje doble alcanza un AUC de 1 y la clase de empuje pendular obtiene un AUC de 0.98. Tanto la clase de empuje clásico como la clase de empuje pendular tienen una forma convexa que se acerca a la esquina superior izquierda del gráfico, lo que denota su buen desempeño. La curva de la clase pendular no posee esta forma convexa tan marcada y tarda un poco en alcanzar 1 en el eje de la tasa de verdaderos positivos, lo que denota que el modelo tiene dificultades para clasificar correctamente los casos positivos al principio.

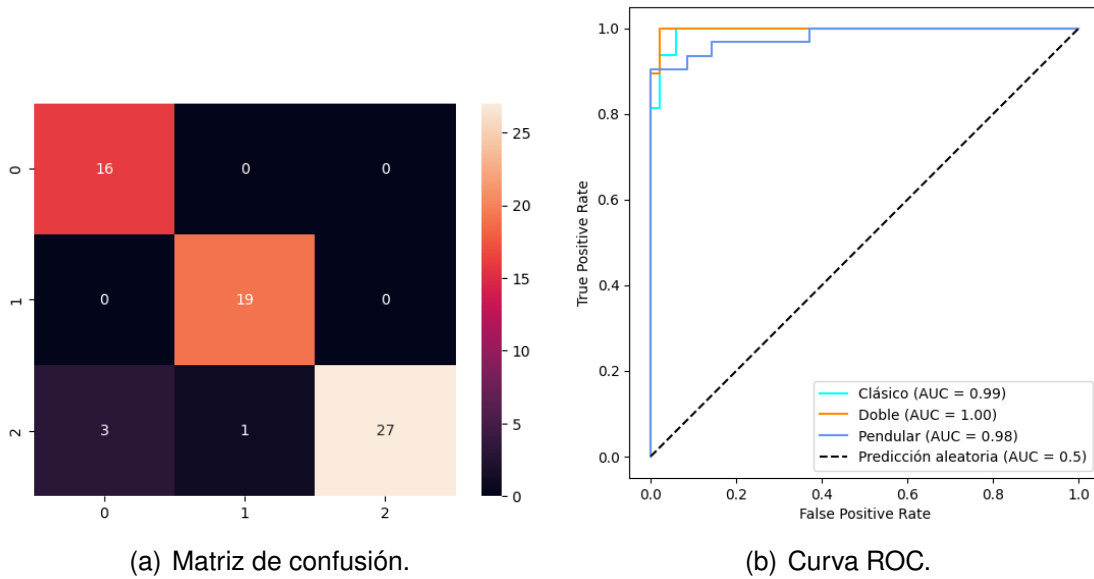


Figura 5.6: Evaluación de SVM.

### 5.1.3.3. Clasificación usando Random Forest (RF)

Como contraste a los métodos utilizados, se presentan los resultados obtenidos al aplicar otro enfoque de clasificación utilizando el algoritmo de RF. Al utilizar los mismos conjuntos de entrenamiento y prueba que para SVM, se logró una exactitud del 100% en el conjunto de entrenamiento y del 92.42% en el conjunto de pruebas, lo cual indica un posible sobreajuste del clasificador. El tiempo de ejecución para clasificación con RF es de 2.349 segundos.

La Tabla 5.6 muestra más métricas del clasificador.

Tabla 5.6: Evaluación de diferentes métricas para RF

Tipos de empuje	Precisión (%)	Sensibilidad (%)	F1-score (%)
Clásico	93.75	93.75	93.75
Doble	90.00	94.74	92.31
Pendular	93.33	90.32	91.80

El empuje clásico muestra un buen rendimiento en todas las métricas, con una precisión, sensibilidad y F1-score del 93.75%. El empuje doble tiene una precisión ligeramente menor, pero una sensibilidad más alta. Esto significa que el modelo tiene una mayor capacidad para detectar correctamente los casos positivos, aunque puede haber una pequeña cantidad de falsos positivos. La clase de empuje pendular muestra

una precisión y F1-score similares a la clase de empuje clásico, pero con una sensibilidad ligeramente menor de 90.32 %. Esto indica que el modelo puede tener dificultades para clasificar correctamente algunos casos positivos para este tipo de empuje en particular. La Figura 5.7 muestra la matriz de confusión y la curva ROC respectivas. En la Figura 5.7(b) se puede apreciar que la clase de empuje clásico tiene un AUC de 0.99, la clase de empuje doble tiene un AUC de 1 y la clase de empuje pendular posee un AUC de 0.98. La curva de la clase pendular no es tan convexa como las otras dos, lo cual puede denotar problemas de desequilibrio entre las clases.

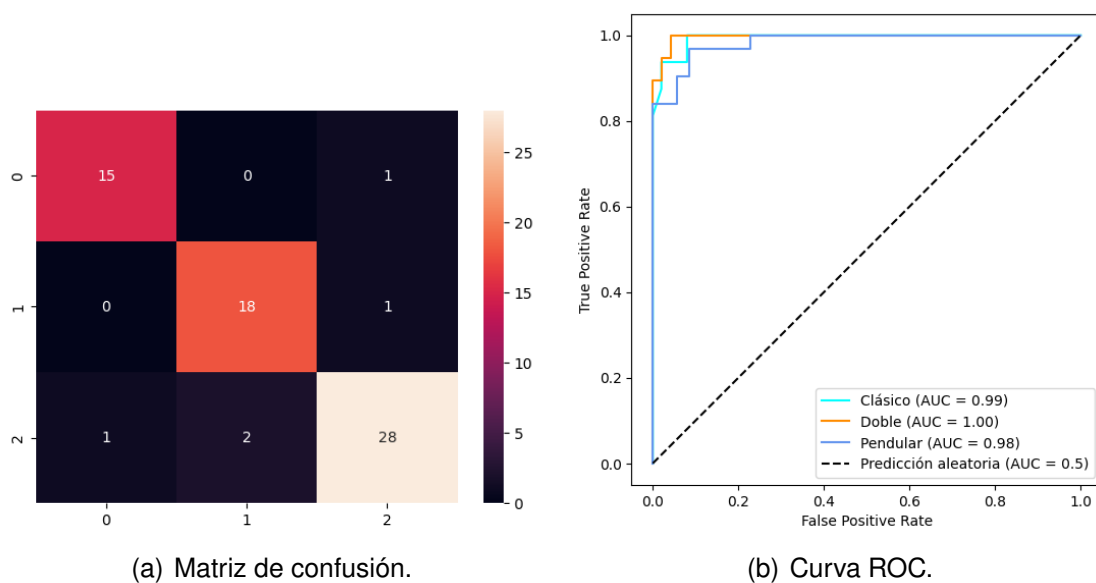


Figura 5.7: Métodos de evaluación.

#### 5.1.3.4. Comparación entre los métodos de clasificación

Los tres sistemas expuestos demuestran una alta exactitud de clasificación en los conjuntos de prueba respectivos, con un 90.72 % para la CNN, 93.94 % para SVM y un 92.42 % para RF. Además, varias métricas adicionales para cada sistema de clasificación también superan el 90 %. Esto se refleja en las matrices de confusión y las curvas ROC, que muestran resultados prometedores en la tarea de clasificación de tipos de empuje de un patinador en recta. Por lo tanto, la elección del un sistema de clasificación estaría determinada por otras características de los sistemas, además del nivel de exactitud, como la facilidad de implementación, la capacidad de aplicar *Data Augmentation*, tiempo de procesamiento, entre otras. Es importante considerar

el tipo de procesamiento realizado por los algoritmos. Por un lado, CNN requiere más tiempo de ejecución debido al entrenamiento de un conjunto de imágenes. Sin embargo, incluye una fase de validación que configura automáticamente los parámetros. Por otro lado, con SVM y RF se tiene que recurrir a otros procesos para encontrar los valores óptimos de los parámetros, como por ejemplo la búsqueda iterativa o aleatoria de valores óptimos. Además, la generación de nuevos datos para aumentar la base de entrenamiento es más sencilla con CNN, ya que es posible realizar transformaciones sobre las imágenes de entrenamiento manteniendo las características distintivas de la figura. Esto, en un aspecto teórico, no crea nueva información, pero ayuda a que el conjunto de entrenamiento tenga características de *BigData* y que el modelo entrenado pueda extraer y aprender características y patrones significativos de mejor manera. En cambio, SVM y RF se ven limitados en la generación de nuevos datos debido a la naturaleza de los vectores de características. Sin embargo, existen técnicas de reducción de características que facilitan el manejo de los datos y permiten extraer la información más relevante en estos casos.

También se debe considerar el tiempo de ejecución de cada algoritmo. El modelo de clasificación de imágenes mediante CNN demora más en ser entrenado en relación a los métodos de clasificación de vectores de características con SVM y RF. Por otra parte, utilizar una CNN permite visualizar de forma más sencilla los cambios que se producen durante el entrenamiento en relación a curvas de pérdida y exactitud. También se debe analizar la existencia de sobreajuste en los modelos obtenidos. Particularmente el caso del clasificador RF muestra una tendencia a un posible sobreajuste, ya que su exactitud en el conjunto de entrenamiento es mayor a la exactitud obtenida en el conjunto de pruebas.

Analizando los datos de las Tablas 5.4, 5.5 y 5.6 se puede observar que los tres sistemas presentan resultados sólidos en las diferentes clases. Sin embargo, hay algunas diferencias en el rendimiento de cada modelo según la métrica y el tipo de empuje. Por ejemplo, SVM muestra una alta sensibilidad, 100%, en la clasificación para las clases de clásico y doble, mientras que RF obtiene una alta precisión en la clase de empuje clásico, 93.75%, cuyo valor es igual a su sensibilidad y consecuentemente a su F1-score. La CNN se mantiene equilibrada en términos de precisión y sensibilidad



en la mayoría de los casos. Analizando los valores de F1-score y las gráficas de curva ROC se puede apreciar que la clase que mejor desempeño muestra en los tres sistemas es el empuje doble. Así mismo se puede ver que CNN clasifica mejor el empuje pendular que el clásico y RF clasifica mejor el empuje clásico que el pendular.

La Tabla 5.7 resume algunas de las características más importantes de los sistemas de clasificación.

Tabla 5.7: Características de los sistemas de clasificación.

Clasificador	Exactitud		Tiempo de ejecución (s)
	entrenamiento (%)	Exactitud Prueba (%)	
CNN	80.97	90.72	173.689
SVM	92.22	93.94	0.692
RF	100.00	92.42	2.349

## Capítulo 6 Conclusiones y recomendaciones

### 6.1. Conclusiones

Se ha realizado un estudio de los modelos biomecánicos relacionados con la técnica del patinaje en línea recta. Se han identificado características distintivas entre la ejecución de tres tipos de empuje por parte de los patinadores, y se han descrito las principales fases del movimiento durante la ejecución de esta técnica. Estos análisis se basan en trabajos y bibliografía especializada que abordan en detalle estos temas. Además, en el Capítulo 3, se han examinado diversos estudios relacionados con los campos de Computer Vision (CV) y Machine Learning (ML) en el ámbito deportivo. Se han tenido en cuenta investigaciones que abordan deportes como el tenis, el tenis de mesa, el béisbol, el baloncesto, entre otros, en los cuales se han empleado herramientas de CV como *OpenPose*, así como métodos de clasificación como Convolutional Neural Network (CNN), Support Vector Machine (SVM) y Random Forest (RF). Se muestra como *OpenPose* proporciona un enfoque no invasivo para generar variables dinámicas relevantes en el estudio y mejoramiento de un deporte en particular. Este enfoque permite a los deportistas ejercer su técnica sin depender de aparatos que limiten su movilidad o elementos que interfieran con la correcta aplicación de una técnica deportiva. Los resultados respaldan la utilidad de *OpenPose* como una herramienta valiosa en la mejora del rendimiento deportivo en conjunto con técnicas de ML para la clasificación y distinción de movimientos.

Se ha logrado un procesamiento efectivo de los datos de un video, transformándolos en otro formato de información, imágenes y vectores de características. Como resultado, se ha creado un conjunto de datos para el entrenamiento, validación y prueba de métodos de clasificación basados en ML. Se han obtenido satisfactoriamente imágenes Skeleton Gait Energy Image (SGEI) que contienen información sobre el movimiento de los patinadores. Además, se ha generado un conjunto de vectores de características, que se forman a partir de variables dinámicas extraídas de las coordenadas de los puntos clave de los patinadores. Los resultados obtenidos en relación a las variables dinámicas que representan la biomecánica del movimiento de un patinador en línea recta, incluyendo distancias, ángulos, tiempos y velocidades, son consistentes tanto con los valores encontrados en la bibliografía de referencia como

con los movimientos de los deportistas observadas durante la ejecución de la técnica. Estos hallazgos proporcionan un respaldo sólido para la evaluación de los resultados y brindan un espacio propicio para la implementación de guías y asesoramiento para los entrenadores de esta disciplina. Los datos obtenidos son de gran relevancia en el estudio y análisis del rendimiento de los patinadores, y ofrecen una base sólida para futuras investigaciones en este campo.

En este estudio, se propuso la implementación de sistemas de clasificación basados en ML para la clasificación de tipos de empuje de patinadores en recta. Se exploraron dos enfoques principales para lograr esta clasificación. El primer enfoque utiliza la clasificación de imágenes SGEI mediante el uso de una CNN. Se logra una exactitud del 90.72 % en la clasificación de datos utilizando la arquitectura Visual Geometry Group (VGG19) como base. Por otro lado, se aplica el enfoque de vectores de características obtenidos a través de un análisis biomecánico. Se utiliza un sistema de SVM y se compararon sus resultados con un algoritmo de RF. El sistema SVM alcanza una exactitud del 93.94 %, mientras que el algoritmo RF logra un 92.42 %. Se presentan diversas métricas de evaluación para los sistemas de clasificación, incluyendo precisión, sensibilidad y F1-score. Además, se muestran resultados gráficos del desempeño de los sistemas, como matrices de confusión y curvas Receiver Operating Characteristic (ROC). También se analizan características de los dos enfoques principales, CNN y SVM, como el aumento de datos y cómo puede ayudar a combatir el sobreajuste. Además, se incluyen tablas con los resultados de las variables biomecánicas obtenidas durante el desarrollo del trabajo, así como imágenes de los ángulos más característicos durante la ejecución de los diferentes tipos de empuje en técnica de recta del patinaje.

Tanto el enfoque de CNN como el de SVM ofrecen resultados prometedores, con altos niveles de exactitud en la clasificación de los datos. Las métricas de evaluación y los análisis realizados proporcionaron una comprensión detallada del rendimiento de los sistemas, y se exploraron estrategias para mejorar su desempeño. Estos hallazgos contribuyen al campo de la clasificación de tipos de empuje en la técnica de recta de patinaje y ofrecen perspectivas valiosas para su aplicación práctica. Aun así, la elección del sistema óptimo dependerá también de otros factores, además de las

métricas obtenidas como la facilidad de implementación, la capacidad de aplicar *Data Augmentation*, tiempo de procesamiento o si es posible implementar métodos complementarios para mejorar la exactitud. Uno de los principales desafíos en los sistemas de clasificación es evitar un excesivo sobreajuste, por lo que sería beneficioso realizar un análisis de cómo generalizan los sistemas al disponer de más datos.

## 6.2. Recomendaciones

La recopilación de datos es un proceso crucial en este trabajo. Mejorar cualquier factor que contribuya a capturar videos de mejor calidad, como el uso de cámaras más avanzadas, permitirá obtener datos más precisos y limpios. Es igualmente importante tener la mayor cantidad posible de datos, por lo que contar con equipos de calidad brinda un mejor control sobre la cantidad de datos utilizables, evitando tener que descartar videos.

En cuanto a la vestimenta de los deportistas, se recomienda utilizar prendas reconocibles y distinguibles, ya que esto facilita la captura de datos en *OpenPose*. Por ejemplo, el uso de licras de color negro completo dificulta la correcta identificación de las partes del cuerpo.

Para procesar los datos, es necesario disponer de una máquina con los recursos adecuados. Esto implica contar con un equipo de gama alta que cuente con una memoria Random Access Memory (RAM) y Video Random Access Memory (VRAM) de alto rendimiento, o bien, optar por servicios de computación en la nube, como ha sido el caso en este trabajo.

## 6.3. Trabajos futuros

Se pueden considerar diferentes enfoques para futuras investigaciones. En primer lugar, sería beneficioso recolectar una mayor cantidad de datos, ya que un sistema de clasificación se beneficia del aumento de información disponible, lo cual permitiría identificar patrones más sólidos y generalizar parámetros de identificación y clasificación de una mejor manera. Con este fin, se buscaría la participación de un número

mayor de participantes.

Además, sería válido explorar la incorporación de ángulos de cámara adicionales y otras medidas complementarias, con el objetivo de lograr una clasificación aún más precisa. Asimismo, sería beneficioso investigar otros patrones y cambios biomecánicos relevantes en el deporte, como las técnicas de salida, en curva y de llegada. Igualmente, sería valioso considerar la posibilidad de prevenir lesiones o identificarlas de manera temprana en los atletas.

Además, se sugiere considerar el uso de un algoritmo de *reinforcement learning* que podría ayudar a mejorar la precisión de la clasificación y permitir adaptaciones en tiempo real de diferentes entornos y condiciones de patinaje de velocidad. Este enfoque podría proporcionar una mayor robustez y generalización del modelo.

## Referencias

- [1] Z. Cao, T. Simon, S.-E. Wei, y Y. Sheikh, "Realtime multi-person 2d pose estimation using part affinity fields," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 1302–1310.
- [2] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, y Y. Sheikh, "Openpose: Realtime multi-person 2d pose estimation using part affinity fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, num. 1, pp. 172–186, 2021.
- [3] G. Hidalgo, Z. Cao, T. Simon, S.-E. Wei, Y. Raaj, H. Joo, y Y. Sheikh, "Openpose: Openpose doc - output," Disponible en [https://cmu-perceptual-computing-lab.github.io/openpose/web/html/doc/md\\_doc\\_02\\_output.html](https://cmu-perceptual-computing-lab.github.io/openpose/web/html/doc/md_doc_02_output.html), 2022, accedido el 20 de marzo de 2023.
- [4] X. Liu y J. Liu, "Gait recognition method of underground coal mine personnel based on densely connected convolution network and stacked convolutional autoencoder," *Entropy*, vol. 22, num. 6, p. 695, 2020.
- [5] Y. Zheng, C. Yang, y A. Merkulov, "Breast cancer screening using convolutional neural network and follow-up digital mammography," in *Computational Imaging III*, vol. 10669. SPIE, 05 2018, p. 4.
- [6] D. Meyer y F. Wien, "Support vector machines," *The Interface to libsvm in package e1071*, vol. 28, p. 20, 2015.
- [7] D. A. Pisner y D. M. Schnyer, "Chapter 6 - support vector machine," in *Machine Learning*, A. Mechelli y S. Vieira, Eds. Academic Press, 2020, pp. 101–121. [En línea]. Disponible: <https://www.sciencedirect.com/science/article/pii/B9780128157398000067>
- [8] Google, "Classification: Roc curve and auc | machine learning | google developers," Disponible en <https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc?hl=en>, 2022, accedido el 28 de marzo de 2023.
- [9] R. Venkatesan y B. Li, *Convolutional neural networks in visual computing: a concise guide*. CRC Press, 2017.

- [10] P. Yeşil, “The importance of physical activity in the prevention and control of cardiovascular diseases,” *Journal of Cardiovascular Nursing*, vol. 3, num. 1, pp. 39–48, 2012.
- [11] H. Ramos, “Vista de Patinaje de velocidad: Revisión sistemática,” *Edu-fisica.com Ciencias aplicadas al deporte*, vol. 11, pp. 143–153, mar 2019. [En línea]. Disponible: <http://revistas.ut.edu.co/index.php/edufisica/article/view/1325/1263>
- [12] P. Madeleine, A. Samani, M. De Zee, y U. Kersting, “Biomechanics of human movement,” *IFMBE Proceedings*, vol. 34 IFMBE, pp. 237–240, 2011. [En línea]. Disponible: [https://www.researchgate.net/publication/226258451\\_Biomechanics\\_of\\_Human\\_Movement](https://www.researchgate.net/publication/226258451_Biomechanics_of_Human_Movement)
- [13] “Kinovea.” [En línea]. Disponible: <https://www.kinovea.org/>
- [14] Y.-C. Li, C.-T. Chang, C.-C. Cheng, y Y.-L. Huang, “Baseball swing pose estimation using openpose,” in *2021 IEEE International Conference on Robotics, Automation and Artificial Intelligence (RAAI)*, 2021, pp. 6–9.
- [15] J. Acero, “Modelo biomecánico biomin-patin. texto-guía diplomado biomin-patin ii&sb,” Liga Santandereana de patinaje e Instituto de Investigaciones & Soluciones Biomecánicas, Bucaramanga, 2022.
- [16] I. Takeda, A. Yamada, y H. Onodera, “Artificial intelligence-assisted motion capture for medical applications: a comparative study between markerless and passive marker motion capture,” *Computer methods in biomechanics and biomedical engineering*, vol. 24, num. 8, p. 864–873, 2021.
- [17] P. Jafarzadeh, P. Virjonen, P. Nevalainen, F. Farahnakian, y J. Heikkonen, “Pose estimation of hurdles athletes using openpose,” in *2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, 2021, pp. 1–6.
- [18] Y. Kong, Z. Wei, y S. Huang, “Automatic analysis of complex athlete techniques in broadcast taekwondo video,” *Multimedia Tools and Applications*, vol. 77, num. 11, pp. 13643–13660, 2018. [En línea]. Disponible: <https://doi.org/10.1007/s11042-017-4979-0>

- [19] P. Zhu y F. Sun, "Sports athletes' performance prediction model based on machine learning algorithm," in *International Conference on Applications and Techniques in Cyber Intelligence ATCI 2019*, J. H. Abawajy, K.-K. R. Choo, R. Islam, Z. Xu, y M. Atiquzzaman, Eds. Cham: Springer International Publishing, 2020, pp. 498–505.
- [20] L. Liu, "Objects detection toward complicated high remote basketball sports by leveraging deep cnn architecture," *Future Generation Computer Systems*, vol. 119, pp. 31–36, 2021. [En línea]. Disponible: <https://www.sciencedirect.com/science/article/pii/S0167739X21000303>
- [21] A. Tarun y A. Nandy, "Human gait classification using deep learning approaches," in *Proceedings of International Conference on Computational Intelligence and Data Engineering*, N. Chaki, J. Pejas, N. Devarakonda, y R. M. Rao Kovvur, Eds. Singapore: Springer Singapore, 2021, pp. 185–199.
- [22] L. Muñoz, D. Yovanovic, T. Bonilla, V. Vargas, y C. Fuentes, *Patinaje de carreras Modelo Curricular Incorporación Fundamentación Perfeccionamiento*, 1998.
- [23] G. van Ingen Schenau, R. De Boer, y G. De Groot, "Biomechanics of speed skating," in *Biomechanics of sport*. CRC press, 2020, pp. 121–167.
- [24] D. J. Ruiz Rivera, "Valoración funcional en patinadores de velocidad de alto nivel: determinación de forma directa, mediante una prueba de campo, de la velocidad aeróbica máxima patinando," 2015.
- [25] C. Lugea, "Algunas Consideraciones sobre Biomecánica, Técnica y el Modelo Técnico en el Patinaje de Velocidad." *Aparece en la revista Roller Training* [http://roller-training.blogspot.com/2014/07/biomecanica-tecnica-y-modelo-tecnico-en\\_21.html](http://roller-training.blogspot.com/2014/07/biomecanica-tecnica-y-modelo-tecnico-en_21.html), pp. 1–68, 2015.
- [26] Y. Tomita, T. Iizuka, K. Irisawa, y S. Imura, "Detection of movement events of long-track speed skating using wearable inertial sensors," *Sensors*, vol. 21, num. 11, p. 3649, 2021.
- [27] A. Stuart y K. C. Cochrane-Snyman, "Strength training and development in com-



- petitive speed skating,” *Strength and Conditioning Journal*, vol. 44, num. 3, pp. 1–10, 2022.
- [28] E. van der Kruk, A. L. Schwab, F. C. van der Helm, y H. E. Veeger, “Getting in shape: Reconstructing three-dimensional long-track speed skating kinematics by comparing several body pose reconstruction techniques,” *Journal of Biomechanics*, vol. 69, pp. 103–112, mar 2018.
- [29] J. Acero y A. Palomino, “Modelo de evaluación y control biomecánico (Biomin-Patín) en el entrenamiento de los patinadores de carreras,” *Memorias EXPOMOTRICIDAD 2009*, num. 11, pp. 247–250, 2010.
- [30] E. P. GUALLASAMIN QUEZADA, “Incidencia de la flexibilidad entre principiantes y experimentados en la fase de empuje y recuperación en el patinaje de velocidad en la escuela de patinaje light speed skaters en el sector de parque bicentenario, cotocollao,” 2020.
- [31] R. Barros y J. Prieto, “Fundamentos teóricos que existen respecto al proceso de entrenamiento de la técnica de patinaje de velocidad en los últimos 10 años en el mundo,” pp. 9–35, 2018. [En línea]. Disponible: [www.ucundinamarca.edu.co](http://www.ucundinamarca.edu.co)
- [32] K. García y J. Bolivar, “Comparacion Cinematica de los ciclos de empuje en patinadores competitivos, DE 11 A 17 años, utilizando una tabla deslizante y la recta en pista,” Ph.D. dissertation, 2011.
- [33] J. Lozada, “Material didactico para la formacion de entrenadoras y entrenadores deportivos ,” 2013. [En línea]. Disponible: [https://www.researchgate.net/publication/294305596\\_PATINAJE\\_MANUAL\\_DIDACTICO](https://www.researchgate.net/publication/294305596_PATINAJE_MANUAL_DIDACTICO)
- [34] A. . Beatriz y M. Jiménez, “Biomechanics and more frequent pathologies in the lower limb in skaters,” 2016.
- [35] J. Stenum, C. Rossi, y R. T. Roemmich, “Two-dimensional video-based analysis of human gait using pose estimation,” *PLoS computational biology*, vol. 17, num. 4, p. e1008935, 2021.
- [36] V. Ramos Arim, C. G. FÁjbrica Barrios, V. Silva Pereyra, y T. Camarot GonzÁjlez, “EvaluaciÃde variables espacio-temporales y transferencia de

- energía mecánica en la marcha de un paciente con ACV: estudio piloto,” *Revista MÃdel Uruguay*, vol. 33, pp. 166 – 186, 03 2017. [En línea]. Disponible: [http://www.scielo.edu.uy/scielo.php?script=sci\\_arttext&pid=S1688-03902017000100166&nrm=iso](http://www.scielo.edu.uy/scielo.php?script=sci_arttext&pid=S1688-03902017000100166&nrm=iso)
- [37] S. Benenaula, M. D. Trelles, L. E. Garza-Castañón, y L. I. Minchala, “Classification of Gait Anomalies By Using Space-Time Parameters Obtained With Pose Estimation,” *International Journal of Innovative Computing, Information and Control*, vol. 18, num. 6, pp. 1913–1927, 2022.
- [38] J. Acero, *Biomecánica integral para patinadores de carrera*, 2010, vol. 21, num. 1. [En línea]. Disponible: <http://journal.um-surabaya.ac.id/index.php/JKM/article/view/2203>
- [39] G. Hidalgo, Z. Cao, T. Simon, S.-E. Wei, Y. Raaj, H. Joo, y Y. Sheikh, “Github - cmu-perceptual-computing-lab/openpose: Openpose: Real-time multi-person keypoint detection library for body, face, hands, and foot estimation,” Disponible en <https://github.com/CMU-Perceptual-Computing-Lab/openpose>, 2022, accedido el 20 de marzo de 2023.
- [40] A. Viswakumar, V. Rajagopalan, T. Ray, y C. Parimi, “Human gait analysis using openpose,” in *2019 Fifth International Conference on Image Information Processing (ICIIP)*, 2019, pp. 310–314.
- [41] E. D’Antonio, J. Taborri, I. Mileti, S. Rossi, y F. Patané, “Validation of a 3d markerless system for gait analysis based on openpose and two rgb webcams,” *IEEE Sensors Journal*, vol. 21, num. 15, pp. 17 064–17 075, 2021.
- [42] M. Nakai, Y. Tsunoda, H. Hayashi, y H. Murakoshi, “Prediction of basketball free throw shooting by openpose,” in *New Frontiers in Artificial Intelligence: JSAI-isAI 2018 Workshops, JURISIN, AI-Biz, SKL, LENLS, IDAA, Yokohama, Japan, November 12–14, 2018, Revised Selected Papers*. Springer, 2019, pp. 435–446.
- [43] W. Chen, Z. Jiang, H. Guo, y X. Ni, “Fall detection based on key points of human-skeleton using openpose,” *Symmetry*, vol. 12, num. 5, p. 744, 2020.

- [44] A. Kececi, A. Yildirak, K. Ozyazici, G. Ayluctarhan, O. Agbulut, y I. Zincir, “Implementation of machine learning algorithms for gait recognition,” *Engineering Science and Technology, an International Journal*, vol. 23, num. 4, pp. 931–937, 2020.
- [45] M. Ota, H. Tateuchi, T. Hashiguchi, y N. Ichihashi, “Verification of validity of gait analysis systems during treadmill walking and running using human pose tracking algorithm,” *Gait & Posture*, vol. 85, pp. 290–297, 2021.
- [46] S. W. Sun, T. C. Mou, C. C. Fang, P. C. Chang, K. L. Hua, y H. C. Shih, “Baseball player behavior classification system using long short-term memory with multimodal features,” *Sensors (Switzerland)*, vol. 19, num. 6, pp. 1–17, 2019.
- [47] E. Madenci, A. Barut, y M. Dorduncu, *Interpolation, Regression, and Smoothing*. Cham: Springer International Publishing, 2019, pp. 57–90. [En línea]. Disponible: [https://doi.org/10.1007/978-3-030-02647-9\\_4](https://doi.org/10.1007/978-3-030-02647-9_4)
- [48] MathWorks, “Suavizado de datos y detección de valores atípicos - matlab amp; simulink - mathworks españa,” Disponible en [https://es.mathworks.com/help/matlab/data\\_analysis/data-smoothing-and-outlier-detection.html](https://es.mathworks.com/help/matlab/data_analysis/data-smoothing-and-outlier-detection.html), 2022, accedido el 21 de marzo de 2023.
- [49] —, “Hann (hanning) window - matlab hann,” Disponible en <https://www.mathworks.com/help/signal/ref/hann.html>, 2022, accedido el 21 de marzo de 2023.
- [50] C. Drentea, *Modern Communications Receiver Design and Technology*, ser. Artech House intelligence and information operations series. Artech House, 2010. [En línea]. Disponible: <https://books.google.com.ec/books?id=9juUwbKP-58C>
- [51] L. Tan y J. Jiang, “Chapter 7 - finite impulse response filter design,” in *Digital Signal Processing (Third Edition)*, third edition ed., L. Tan y J. Jiang, Eds. Academic Press, 2019, pp. 229–313. [En línea]. Disponible: <https://www.sciencedirect.com/science/article/pii/B9780128150719000075>
- [52] J. Loureiro y P. L. Correia, “Using a skeleton gait energy image for pathological

- gait classification,” in *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*, 2020, pp. 503–507.
- [53] J. Luo, J. Zhang, C. Zi, Y. Niu, H. Tian, y C. Xiu, “Gait recognition using gei and afdei,” *International Journal of Optics*, vol. 2015, 2015.
- [54] S. C. Bakchy, M. Islam, M. R. Mahmud, F. Imran y otros, “Human gait analysis using gait energy image,” *arXiv preprint arXiv:2203.09549*, 2022.
- [55] E. Newman, *Matemáticas e imaginación*. CONACULTA, Direccion General de Publicaciones, 2007. [En línea]. Disponible: <https://books.google.com.ec/books?id=zdBHMHV3m5YC>
- [56] M. Peters y R. González, *Álgebra y trigonometría*. Reverte, 2022. [En línea]. Disponible: <https://books.google.com.ec/books?id=H39YEAAAQBAJ>
- [57] C. Aggarwal, *Data Classification: Algorithms and Applications*, ser. Chapman & Hall/CRC Data Mining and Knowledge Discovery Series. Taylor & Francis, 2014. [En línea]. Disponible: <https://books.google.com.ec/books?id=gJhBBAAAQBAJ>
- [58] H. H. Aghdam y E. J. Heravi, “Guide to convolutional neural networks,” *New York, NY: Springer*, vol. 10, num. 978-973, pp. 102–110, 2017.
- [59] L. Lu, Y. Zheng, G. Carneiro, y L. Yang, “Deep learning and convolutional neural networks for medical image computing,” *Advances in computer vision and pattern recognition*, vol. 10, p. 38, 2017.
- [60] K. Simonyan y A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1409.1556*, 2014.
- [61] M. Bansal, M. Kumar, M. Sachdeva, y A. Mittal, “Transfer learning for image classification using vgg19: Caltech-101 image data set,” *Journal of ambient intelligence and humanized computing*, pp. 1–12, 2021.
- [62] S. Suthaharan y S. Suthaharan, “Support vector machine,” *Machine learning models and algorithms for big data classification: thinking with examples for effective learning*, pp. 207–235, 2016.

- [63] S. Haghghi, M. Jasemi, S. Hessabi, y A. Zolanvari, “Pycm: Multiclass confusion matrix library in python,” *Journal of Open Source Software*, vol. 3, num. 25, p. 729, 2018.
- [64] D. K. Sharma, M. Chatterjee, G. Kaur, y S. Vavilala, “3 - deep learning applications for disease diagnosis,” in *Deep Learning for Medical Applications with Unique Data*, D. Gupta, U. Kose, A. Khanna, y V. E. Balas, Eds. Academic Press, 2022, pp. 31–51. [En línea]. Disponible: <https://www.sciencedirect.com/science/article/pii/B9780128241455000058>
- [65] A. Meyer-Baese y V. Schmid, “Chapter 6 - statistical and syntactic pattern recognition,” in *Pattern Recognition and Signal Analysis in Medical Imaging (Second Edition)*, second edition ed., A. Meyer-Baese y V. Schmid, Eds. Oxford: Academic Press, 2014, pp. 151–196. [En línea]. Disponible: <https://www.sciencedirect.com/science/article/pii/B9780124095458000066>
- [66] A. Tiwari, “Chapter 2 - supervised learning: From theory to applications,” in *Artificial Intelligence and Machine Learning for EDGE Computing*, R. Pandey, S. K. Khatri, N. kumar Singh, y P. Verma, Eds. Academic Press, 2022, pp. 23–32. [En línea]. Disponible: <https://www.sciencedirect.com/science/article/pii/B9780128240540000265>
- [67] P. Pareek y A. Thakkar, “A survey on video-based human action recognition: recent updates, datasets, challenges, and applications,” *Artificial Intelligence Review*, vol. 54, num. 3, pp. 2259–2322, 2021. [En línea]. Disponible: <https://doi.org/10.1007/s10462-020-09904-8>
- [68] K. Host y M. Ivašić-Kos, “An overview of human action recognition in sports based on computer vision,” *Heliyon*, vol. 8, num. 6, p. e09633, 2022. [En línea]. Disponible: <https://www.sciencedirect.com/science/article/pii/S2405844022009215>
- [69] Y. Kong y Y. Fu, “Human action recognition and prediction: A survey,” 2022.
- [70] M. Ivašić-Kos, K. Host, y M. Pobar, *Application of Deep Learning Methods for Detection and Tracking of Players*, 02 2021.

- [71] y. Wang, W. Fang, J. Ma, X. Li, y A. Zhong, *Automatic Badminton Action Recognition Using CNN with Adaptive Feature Extraction on Sensor Data*, 07 2019, pp. 131–143.
- [72] K. Joshi, V. Tripathi, C. Bose, y C. Bhardwaj, “Robust sports image classification using inceptionv3 and neural networks,” *Procedia Computer Science*, vol. 167, pp. 2374–2381, 01 2020.
- [73] F. Angelini, Z. Fu, Y. Long, L. Shao, y S. M. Naqvi, “2d pose-based real-time human action recognition with occlusion-handling,” *IEEE Transactions on Multimedia*, vol. 22, num. 6, pp. 1433–1446, 2020.
- [74] N. U. R. Malik, S. Bakar, U. U. Sheikh, A. Channa, y N. Popescu, “Cascading pose features with cnn-lstm for multiview human action recognition,” *Signals*, vol. 4, 01 2023.
- [75] C. Sawant, “Human activity recognition with openpose and long short-term memory on real time images,” 2020.
- [76] F. M. Noori, B. Wallace, M. Z. Uddin, y J. Torresen, “A robust human activity recognition approach using openpose, motion features, and deep recurrent neural network,” in *Image Analysis*, M. Felsberg, P.-E. Forss’en, I.-M. Sintorn, y J. Unger, Eds. Cham: Springer International Publishing, 2019, pp. 299–310.
- [77] B. Abirami, T. Subashini, y V. Mahavaishnavi, “Automatic age-group estimation from gait energy images,” *Materials Today: Proceedings*, vol. 33, pp. 4646–4649, 2020, international Conference on Nanotechnology: Ideas, Innovation and Industries. [En línea]. Disponible: <https://www.sciencedirect.com/science/article/pii/S2214785320361708>
- [78] A. Sampath Dakshina Murthy, T. Karthikeyan, y R. Vinoth Kanna, “Gait-based person fall prediction using deep learning approach,” *Soft Computing*, vol. 26, num. 23, pp. 12 933–12 941, 2022. [En línea]. Disponible: <https://doi.org/10.1007/s00500-021-06125-1>
- [79] V. Podgorelec, Pečnik, y G. Vrbančič, “Classification of similar sports images using convolutional neural network with hyper-parameter optimization,”

- Applied Sciences*, vol. 10, num. 23, 2020. [En línea]. Disponible: <https://www.mdpi.com/2076-3417/10/23/8494>
- [80] O. C. Kurban, N. Calik, y T. Yildirim, “Human and action recognition using adaptive energy images,” *Pattern Recognition*, vol. 127, p. 108621, 2022. [En línea]. Disponible: <https://www.sciencedirect.com/science/article/pii/S0031320322001029>
- [81] S. S. Tabrizi, S. Pashazadeh, y V. Javani, “Comparative study of table tennis forehand strokes classification using deep learning and svm,” *IEEE Sensors Journal*, vol. 20, num. 22, pp. 13 552–13 561, 2020.
- [82] M. Ivašić-Kos y M. Pobar, “Building a labeled dataset for recognition of handball actions using mask r-cnn and stips,” 11 2018, pp. 1–6.
- [83] S. J. Benenaula Armijos y M. D. Trelles Peralta, “Generación de la biomecánica del movimiento de extremidades inferiores.” [En línea]. Disponible: <http://dspace.ucuenca.edu.ec/handle/123456789/36081>
- [84] M. Skublewska-Paszkowska y P. Powroznik, “Temporal pattern attention for multivariate time series of tennis strokes classification,” *Sensors*, vol. 23, num. 5, 2023. [En línea]. Disponible: <https://www.mdpi.com/1424-8220/23/5/2422>
- [85] G. C. Domínguez, E. F. Álvarez, A. T. Córdoba, y D. G. Reina, “A comparative study of machine learning and deep learning algorithms for padel tennis shot classification,” *Soft Computing*, Feb 2023. [En línea]. Disponible: <https://doi.org/10.1007/s00500-023-07874-x>
- [86] J. Jang, A. Ankit, J. Kim, Y. J. Jang, H. Y. Kim, J. H. Kim, y S. Xiong, “A unified deep-learning model for classifying the cross-country skiing techniques using wearable gyroscope sensors,” *Sensors*, vol. 18, num. 11, 2018. [En línea]. Disponible: <https://www.mdpi.com/1424-8220/18/11/3819>
- [87] K. V. Kumar, A. Abraham, y S. Elias, “Quantitative analysis of athlete performance in artistic skating using imu, and machine learning algorithms,” *Design Engineering (Toronto)*, vol. 2021, pp. 11 236– 11 252, 12 2021.

- [88] O. M. H. Rindal, T. M. Seeberg, J. Tjønnås, P. Haugnes, y Sandbakk, “Automatic classification of sub-techniques in classical cross-country skiing using a machine learning algorithm on micro-sensor data,” *Sensors*, vol. 18, num. 1, 2018. [En línea]. Disponible: <https://www.mdpi.com/1424-8220/18/1/75>
- [89] C. Sun y D. Ma, “Svm-based global vision system of sports competition and action recognition,” *Journal of Intelligent Fuzzy Systems*, vol. 40, num. 2, pp. 2265–2276, 2021.
- [90] C. de Investigación y Desarrollo de la Educación y la Academia (CEDIA), “Supercomputador,” <https://cedia.edu.ec/servicio/supercomputador/>, Mayo 2023, accessed: Mayo 11, 2023.
- [91] B. Ulloa y X. Albornoz, “Biomecanicapatinaje,” <https://github.com/fabianastudillo/BiomecanicaPatinaje>, Fecha de acceso: 2023-06-09.
- [92] “ImageNet,” <https://www.image-net.org/download.php>, Mayo 2023, acceso: 5 de mayo de 2023.



## Apéndice A *Openpose*

### A.1. Instalación

A continuación se muestra un *script* de *bash* con los comandos necesarios para la instalación de *Openpose* en un sistema operativo basado en Linux.

```
#!/bin/bash

#Actualizacion de paquetes
apt-get update
apt-get upgrade

#Instalacion de dependencias
apt-get install git
apt-get install wget
apt-get install python3-opencv
apt-get install libopencv-dev
apt-get install libboost-all-dev
apt-get install ffmpeg
apt-get install nvidia-cuda-toolkit
apt-get -qq install -y libatlas-base-dev libprotobuf-dev
    libleveldb-dev libsnappy-dev libhdf5-serial-dev protobuf-
    compiler libgflags-dev libgoogle-glog-dev liblmdb-dev opencv-
    -headers ocl-icd-opencv-dev libviennacl-dev
apt install python3-pip
pip install opencv-python
pip install matplotlib

#Instalacion de OpenPose
wget -q https://cmake.org/files/v3.21/cmake-3.21.0-rc3-linux-
    x86_64.tar.gz
tar xfz cmake-3.21.0-rc3-linux-x86_64.tar.gz --strip-
    components=1 -C /usr/local
```

```
git clone -q --depth 1 https://github.com/CMU-Perceptual-Computing-Lab/openpose.git
sed -i "s/execute_process(COMMAND_git_checkout_master_WORKING_DIRECTORY_${_CMAKE_SOURCE_DIR}\3rdparty\caffe)/execute_process(COMMAND_git_checkout_f019d0dfe86f49d1140961f8c7dec22130c83154_WORKING_DIRECTORY_${CMAKE_SOURCE_DIR}\3rdparty\caffe)/g" openpose/CMakeLists.txt
wget -c http://developer.download.nvidia.com/compute/redist/cudnn/v5.1/cudnn-8.0-linux-x64-v5.1.tgz
#tar -xzf cudnn-8.0-linux-x64-v5.1.tgz -C /usr/local
cd openpose && rm -rf build || true && mkdir build && cd build
&& cmake -DBUILD_PYTHON=ON .. && make -j $(nproc)
```

## A.2. Ejecución

A continuación se muestra la instrucción con la cual se ejecuta el programa de *OpenPose*.

```
./build/examples/openpose/openpose.bin --video $FILE_IN --display 0 --number_people_max=1 --net_resolution "1312x736" --render_pose=0 --scale_number 4 --scale_gap 0.25 --write_json $FILE_OUT
```

El parámetro *video* especifica que la entrada es un video, *display* indica si el programa debe mostrar los resultados durante la ejecución, *number\_people\_max* indica el número de personas a detectar, *net\_resolution* es la resolución a la que se transforman los fotogramas del video, *render\_pose* habilita o deshabilita la sobreposición de pose sobre la imagen original y *write\_json* indica que la salida será una carpeta con archivos JSON correspondientes a cada fotograma.

## Apéndice B Valores incluidos en los vectores de características

### B.1. Máximos, mínimos y valor medio de los ángulos (promedio $\pm$ desviación estándar)

Tabla B.1: Ángulos en vista frontal completos

Características	Clásico (°)	Doble (°)	Pendular (°)
Max exterior pie D	119.75 $\pm$ 8.99	121.76 $\pm$ 4.49	126.45 $\pm$ 6.35
Max exterior pie I	118 $\pm$ 8.57	114.91 $\pm$ 5.79	114.93 $\pm$ 5.32
Min exterior pie D	46.25 $\pm$ 7.04	45.36 $\pm$ 6.86	49.34 $\pm$ 6.29
Min exterior pie I	39.67 $\pm$ 8.77	38.9 $\pm$ 5.72	41.56 $\pm$ 7.18
Media exterior pie D	84.71 $\pm$ 6.53	89.04 $\pm$ 4.7	88.91 $\pm$ 4.91
Media exterior pie I	79.56 $\pm$ 4.53	82.82 $\pm$ 4.63	79.81 $\pm$ 5.25
Max interior pie D	133.75 $\pm$ 7.04	134.64 $\pm$ 6.86	130.66 $\pm$ 6.29
Max interior pie I	140.33 $\pm$ 8.77	141.1 $\pm$ 5.72	138.44 $\pm$ 7.18
Min interior pie D	60.25 $\pm$ 8.99	58.24 $\pm$ 4.49	53.55 $\pm$ 6.35
Min interior pie I	62 $\pm$ 8.57	65.09 $\pm$ 5.79	65.07 $\pm$ 5.32
Media interior pie D	95.29 $\pm$ 6.53	90.96 $\pm$ 4.7	91.09 $\pm$ 4.91
Media interior pie I	100.44 $\pm$ 4.53	97.18 $\pm$ 4.63	100.19 $\pm$ 5.25
Max empuje interno D	19.27 $\pm$ 4.26	21.59 $\pm$ 3.73	20.1 $\pm$ 3.17
Max empuje interno I	14.57 $\pm$ 3.64	17.5 $\pm$ 3.47	16.33 $\pm$ 3.31
Min empuje interno D	0.34 $\pm$ 0.3	0.44 $\pm$ 0.36	0.41 $\pm$ 0.33
Min empuje interno I	0.25 $\pm$ 0.24	0.36 $\pm$ 0.32	0.34 $\pm$ 0.3
Media empuje interno D	10.71 $\pm$ 1.86	12.02 $\pm$ 2.03	12.92 $\pm$ 2.17
Media empuje interno I	7.8 $\pm$ 1.86	8.99 $\pm$ 1.92	9.57 $\pm$ 2.4
Max empuje externo D	35.19 $\pm$ 3.53	35.52 $\pm$ 2.37	36.62 $\pm$ 4.24
Max empuje externo I	34.26 $\pm$ 4.11	36.58 $\pm$ 2.54	37.74 $\pm$ 4.15
Min empuje externo D	0.4 $\pm$ 0.32	0.43 $\pm$ 0.33	0.39 $\pm$ 0.32
Min empuje externo I	0.32 $\pm$ 0.27	0.34 $\pm$ 0.34	0.29 $\pm$ 0.25
Media empuje externo D	18.72 $\pm$ 2.47	19.2 $\pm$ 1.74	20.31 $\pm$ 3.17
Media empuje externo I	17.03 $\pm$ 2.5	19.71 $\pm$ 1.62	19.49 $\pm$ 2.69
Max tobillo D	100.96 $\pm$ 7.32	100.57 $\pm$ 6.39	109.27 $\pm$ 14.35
Max tobillo I	100.23 $\pm$ 7.2	102.47 $\pm$ 7.85	106.96 $\pm$ 16.27
Min tobillo D	57.06 $\pm$ 20.29	70.65 $\pm$ 6.01	59.15 $\pm$ 18.35
Min tobillo I	55.44 $\pm$ 19.39	70.49 $\pm$ 8.53	62.19 $\pm$ 17.48
Media tobillo D	80.36 $\pm$ 4.27	82.04 $\pm$ 3.32	78.88 $\pm$ 4.79
Media tobillo I	81.51 $\pm$ 4.22	84.15 $\pm$ 3.11	81.74 $\pm$ 4.4
Max entropierna	73.46 $\pm$ 10.27	74.76 $\pm$ 8.22	76.05 $\pm$ 9.28
Min entropierna	21.29 $\pm$ 9.45	20.56 $\pm$ 7.39	25.06 $\pm$ 6.21
Media entropierna	47.57 $\pm$ 6.24	46.58 $\pm$ 5.88	48.49 $\pm$ 5.85

Tabla B.2: Ángulos en vista sagital completos (promedio  $\pm$  desviación estándar)

<b>Características</b>	<b>Clásico (°)</b>	<b>Doble (°)</b>	<b>Pendular (°)</b>
Max Tronco	30.57 $\pm$ 5.87	27.1 $\pm$ 3.96	29.78 $\pm$ 5.02
Min Tronco	13.99 $\pm$ 5.26	10.1 $\pm$ 3.47	17.45 $\pm$ 4.52
Media Tronco	21.25 $\pm$ 4.19	17.87 $\pm$ 2.74	23.45 $\pm$ 4.64
Max Cadera D	117.98 $\pm$ 11.65	103.82 $\pm$ 9.9	117.8 $\pm$ 10.32
Max Cadera I	123.7 $\pm$ 10.09	118.47 $\pm$ 10.82	124.48 $\pm$ 13.59
Min Cadera D	63.72 $\pm$ 6.58	62.27 $\pm$ 4.97	66.61 $\pm$ 6.77
Min Cadera I	57.99 $\pm$ 6.08	57.29 $\pm$ 7.16	60.14 $\pm$ 7.69
Media Cadera D	86.63 $\pm$ 4.52	80.13 $\pm$ 3.93	87.42 $\pm$ 6.3
Media Cadera I	83.65 $\pm$ 6.24	81.15 $\pm$ 5.45	85.09 $\pm$ 8.73
Max Rodilla D	150.82 $\pm$ 14.22	143.66 $\pm$ 8.14	145.31 $\pm$ 13.93
Max Rodilla I	153.6 $\pm$ 11.75	151.47 $\pm$ 7.81	152.45 $\pm$ 12.69
Min Rodilla D	86.41 $\pm$ 12.13	95.52 $\pm$ 8.4	87 $\pm$ 11.32
Min Rodilla I	90.09 $\pm$ 12.27	91.61 $\pm$ 10.3	81.92 $\pm$ 13.15
Media Rodilla D	119.01 $\pm$ 7.1	118.66 $\pm$ 4.42	116.19 $\pm$ 7.31
Media Rodilla I	118.18 $\pm$ 5.71	120.3 $\pm$ 5.72	116.33 $\pm$ 6.53
Max Tobillo D	135.95 $\pm$ 24.81	126.25 $\pm$ 13.56	132.17 $\pm$ 26.7
Max Tobillo I	135.11 $\pm$ 25.71	133.52 $\pm$ 15.33	130.23 $\pm$ 20.26
Min Tobillo D	76.93 $\pm$ 13.81	82.15 $\pm$ 9.21	77.94 $\pm$ 10.56
Min Tobillo I	78.34 $\pm$ 9.16	81.69 $\pm$ 7.41	76.92 $\pm$ 11.2
Media Tobillo D	102.23 $\pm$ 8.62	101.2 $\pm$ 4.05	99.88 $\pm$ 5.84
Media Tobillo I	97.58 $\pm$ 5.06	101.1 $\pm$ 4.59	98.35 $\pm$ 6.41
Max Canilla D	83.72 $\pm$ 8.74	80.55 $\pm$ 4.38	79.99 $\pm$ 6.6
Max Canilla I	78.38 $\pm$ 7.66	80.27 $\pm$ 5.59	79.98 $\pm$ 6
Min Canilla D	6.51 $\pm$ 14.57	19.6 $\pm$ 7.78	2.03 $\pm$ 10.23
Min Canilla I	12.37 $\pm$ 10.58	18.17 $\pm$ 9.39	5.07 $\pm$ 10.34
Media Canilla D	53.8 $\pm$ 6.9	57.74 $\pm$ 3.13	52.25 $\pm$ 4.99
Media Canilla I	54.85 $\pm$ 5.9	55.65 $\pm$ 6.09	54.14 $\pm$ 6.78
Max Entrepiera	75.98 $\pm$ 12.36	73.21 $\pm$ 12.72	77.48 $\pm$ 13.19
Min Entrepiera	0.28 $\pm$ 0.27	0.28 $\pm$ 0.3	0.28 $\pm$ 0.31
Media Entrepiera	31.29 $\pm$ 7.46	31.18 $\pm$ 6.91	33.1 $\pm$ 7.34
Max Codo D	265.6 $\pm$ 86.69	297.5 $\pm$ 71.13	232.89 $\pm$ 87.89
Max Codo I	259.1 $\pm$ 74.06	236.33 $\pm$ 69.92	257.89 $\pm$ 71.4
Min Codo D	27.28 $\pm$ 23.84	24.22 $\pm$ 24.52	29 $\pm$ 21.56
Min Codo I	38.48 $\pm$ 26.87	32.63 $\pm$ 18.3	42.77 $\pm$ 29.34
Media Codo D	125.94 $\pm$ 18.84	132.01 $\pm$ 17.05	116.27 $\pm$ 25.44
Media Codo I	141.96 $\pm$ 14.82	129.6 $\pm$ 21.5	141.2 $\pm$ 14.75
Max Hombro D	104.93 $\pm$ 21.81	96.08 $\pm$ 15.8	96.16 $\pm$ 14.72
Max Hombro I	92.44 $\pm$ 15.5	88.08 $\pm$ 15.98	85.15 $\pm$ 18.33
Min Hombro D	-21.59 $\pm$ 44.89	-50.52 $\pm$ 39.86	-6.12 $\pm$ 46.75
Min Hombro I	-36.4 $\pm$ 22.06	-50.27 $\pm$ 26.03	-46.61 $\pm$ 26.84
Media Hombro D	51.57 $\pm$ 12.15	41.24 $\pm$ 15.23	53.07 $\pm$ 17.62
Media Hombro I	33.83 $\pm$ 10.14	27.8 $\pm$ 12.82	25.17 $\pm$ 17.43

## Anexos

### Anexo A Permiso

#### A.1. Permiso Club deportivo

Señor

Ingeniero

Alcides Fabian Araujo Pacheco

**Director de la carrera de Ingeniería Electrónica y Telecomunicaciones**

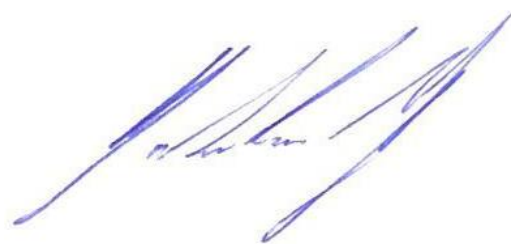
De mis consideraciones:

Reciba un cordial saludo por parte de la dirección del Club Deportivo Especializado Formativo Gabriela Cavalieri.

Por medio de la presente me dirijo a usted para dar a conocer mi apoyo total a los estudiantes egresados de la carrera de Ingeniería en Electrónica y Telecomunicaciones Ximena Lorena Albornoz Tepán con C.I: 0107270910, y Bryan Sebastián Ulloa Montaleza con C.I 0107354334, con el fin de que realicen los estudios necesarios para su trabajo de titulación relacionado de manera implícita con el deporte de Patinaje de Velocidad, para lo cual se proporciona el permiso y acceso absoluto, en el tiempo indefinido que demore el desarrollo del mismo, a las instalaciones, y equipos que forman parte del Patinódromo Ubicado en Totoracocha además de los deportistas que forman parte del club Deportivo Especializado Formativo Gabriela Cavalieri, requeridos para realizar su trabajo de titulación previo a la obtención de su título.

Por la acogida de la presente, anticipo mi agradecimiento.

Atentamente,



Gabriela Cavalieri Zamora

Directora del club deportivo especializado  
formativo Gabriela Cavalieri

## Anexo B Consentimiento informado

### B.1. Consentimiento informado menor de edad

# Consentimiento Informado

(Menor de edad)

Yo ..... , con C.I ..... en fecha ..... de 2023, en representación de ..... en calidad de .....

Manifiesto que he recibido toda la información necesaria, de forma clara, comprensible y satisfactoria sobre la naturaleza y propósito de los objetivos del proceso a seguir.

Por todo lo que se me ha explicado en forma oral presto mi conformidad/consentimiento para que mi hijo/a participe en la adquisición de video como parte del proyecto de investigación "Análisis comparativo entre sistemas de clasificación basados en técnicas de *machine learning* para catalogar tipos de empuje de un patinador de velocidad." Realizado por investigadores de la Universidad de Cuenca, que servirá para evaluar patrones en la técnica de cada patinador; del cual declaro conocer sus finalidades y forma de realización del procedimiento, que implica:

- Patinar frente a 3 cámaras ubicadas sagital y frontalmente.
- Observación directa de la técnica y los tipos de empuje.
- Procesamiento de video mediante visión artificial.
- Duración del estudio 1 hora y 30 minutos.

Se guardará la confidencialidad de la información.

.....

Firma del tutor

.....

Firma del deportista

**B.2. Consentimiento informado mayor de edad**

## Consentimiento Informado

Yo ..... , con C.I ..... en fecha ..... de 2023.

Manifiesto que he recibido toda la información necesaria, de forma clara, comprensible y satisfactoria sobre la naturaleza y propósito de los objetivos del proceso a seguir.

Por todo lo que se me ha explicado en forma oral acepto participar en la adquisición de video como parte del proyecto de investigación "Análisis comparativo entre sistemas de clasificación basados en técnicas de *machine learning* para catalogar tipos de empuje de un patinador de velocidad." Realizado por investigadores de la Universidad de Cuenca, que servirá para evaluar patrones en la técnica de cada patinador; del cual declaro conocer sus finalidades y forma de realización del procedimiento, que implica:

- Patinar frente a 3 cámaras ubicadas sagital y frontalmente.
- Observación directa de la técnica y los tipos de empuje.
- Procesamiento de video mediante visión artificial.
- Duración del estudio 1 hora y 30 minutos.

Se guardará la confidencialidad de la información.

.....

Firma

Anexo C Registro de firmas

C.1. Menores de edad

Nombre del deportista	Nombre del representante	Cedula	Firma del representante	Firma del deportista
Agustín Cabrera	Marina Molina	0999362126		
Ámbar León	Marcelo León	0102944682		
Ariana Pesantez	Pablo Pesantez	0103874103		
Camila Alejandra Narváez	Clara Samaniego	1400344471		
Joaquín Loyola	Gabriela Cavalieri	0103720470		
Jorge Arévalo	Jorge Arévalo	0102361219		
Juan José Loyola	Gabriela Cavalieri	0103720470		
Juliana Larriva	María Verónica Ortiz	0103798658		
María Paula Escandón	Lorena Astudillo	0103473815		
María Paz Guerrero	Patricia Toral	0104285437		
Martina Isabel Rosero	Juan Sebastián Rosero	1801880959		
Mateo Astudillo	Fabian Astudillo	0103907036		
Nicole Segarra	Judith Guillen	0101500817		
Pamela García	Franklin García	0102622438		
Victoria Larriva	Cristina Carpio	0103289402		



## C.2. Mayores de edad

Nombre del deportista	Cedula	Firma del deportista
David Sarmiento	0105002522	
Emilia Escandón	0150038457	
Micaela Matute	0105816920	
Valentina Vélez	0100643674	Valentina Vélez