



UNIVERSIDAD DE VALLADOLID

FACULTAD DE MEDICINA

ESCUELA DE INGENIERÍAS INDUSTRIALES

TRABAJO DE FIN DE GRADO

GRADO EN INGENIERÍA BIOMÉDICA

**Segmentación y análisis de patrones de marcha en
paciente pediátrico con parálisis cerebral
empleando sensores inerciales, visión artificial y
aprendizaje profundo**

Autor/a:

D.^a Ángela Romo Marín

Tutor/a:

D. Mario Martínez Zarzuela

D.^a Cristina Simón Martínez

Valladolid, 13 de julio de 2022

TÍTULO: Segmentación y análisis de la marcha en paciente pediátrico con parálisis cerebral empleando sensores inerciales, visión artificial y aprendizaje profundo

AUTOR/A: **D.ª Ángela Romo Marín**

TUTOR/A: **D. Mario Martínez Zarzuela**
D.ª Cristina Simón Martínez

DEPARTAMENTO: **Departamento de Teoría de la Señal y Comunicaciones e Ingeniería Telemática (TSCIT)**

TRIBUNAL

PRESIDENTE: **D. Carlos Gómez Peña**

SECRETARIO: **D. Mario Martínez Zarzuela**

VOCAL: **D. Daniel Álvarez González**

SUPLENTE 1: **D. Alberto Mansilla Gallo**

SUPLENTE 2: **D. Jesús Poza Crespo**

FECHA:

CALIFICACIÓN:

Agradecimientos

Me gustaría comenzar agradeciendo a Mario Martínez Zarzuela y a Cristina Simón Martínez, mis tutores de TFG, por haberme dado la oportunidad de sumarme a este proyecto, por la ayuda recibida y por la paciencia mostrada ante cada una de mis dudas.

También, agradecer a los profesionales médicos y familias, los cuales emplearon parte de su tiempo libre en acudir a las grabaciones, y han hecho posible la adquisición de la base de datos empleada en este trabajo.

A todos mis amigos, tanto de la universidad como del instituto, por el apoyo incondicional que me han brindado a lo largo de estos años y por haber creído siempre en mí y en que lograría este proyecto.

Por último, a mi familia, en especial a mis padres y a mi hermana, gracias por haberme dedicado todo vuestro tiempo, por confiar ciegamente en mí y por haber permanecido a mi lado, tanto en los buenos como en los malos momentos.

Resumen

La parálisis cerebral (PC) ha sido reconocida como la discapacidad motora más común en la infancia (Bekteshi et al., 2023). La afectación de la marcha en niños y niñas que sufren de esta condición constituye una de las principales limitaciones a las que se ven sometidos en su día a día (Nowicki, 2020). Es por ello que el análisis clínico de la marcha (*Clinical Gait Analysis*, CGA) ha pasado a ser una práctica habitual en los centros sanitarios. En él se miden y analizan múltiples parámetros, con el objetivo de identificar y entender tanto las distintas lesiones del movimiento como sus causas (Armand et al., 2016).

Como complemento a la inspección visual médica ha surgido el análisis instrumentado, el cual se apoya en el uso de dispositivos electrónicos con el fin de brindar un estudio más completo e individualizado a cada uno de los sujetos (Nowicki, 2020). Actualmente, son muchas las opciones de dispositivos de esta clase ofertadas en el mercado. La visión artificial con múltiples cámaras constituye la técnica *gold standard* en el estudio del movimiento (Tunca et al., 2017). Sin embargo, su elevado precio y su necesidad de amplios espacios y formación profesional hacen imposible la generalización del análisis instrumentado de la marcha en el ámbito médico (Gao et al., 2022).

Por ello, durante este estudio se realizó un CGA a partir de unidades de medición inercial (IMU) y sistemas de vídeo. Con ello se pretende demostrar el potencial de ambas técnicas, mucho más económicas que la visión artificial, las cuales podrían hacer posible la extensión de este tipo de práctica en el campo hospitalario. Así pues, se adquirió una base de datos de 25 sujetos, de los cuales 17 eran controles y 8 pacientes pediátricos con PC. A mayores, se elaboró un modelo de *deep learning* capaz de segmentar automáticamente la marcha en sus respectivas fases.

Los resultados obtenidos durante los distintos entrenamientos mostraron que, en sujetos sin afectación de la marcha, con 3-4 sensores inerciales y una arquitectura de red sencilla es posible hacer una segmentación y análisis de gran calidad. Prueba de ello son las métricas de rendimiento conseguidas, entre las que cabe destacar un *accuracy* comprendido entre 0.97 y 1. Por otro lado, también se pudo comprobar cómo, en pacientes con parálisis cerebral, es necesario entrenar modelos que tengan en cuenta todas las posibles desviaciones de la marcha, puesto que no es posible emplear la misma red en ambos grupos de sujetos. Por último, tras el conjunto de pruebas realizadas, quedó demostrada la efectividad de ambas técnicas de medición. En consecuencia, gracias a las mismas, es posible conseguir, de una manera más económica, un estudio objetivo e individualizado de la marcha de cada paciente.

Palabras clave

Parálisis cerebral, análisis instrumentado de la marcha, unidades de medición inercial, sistemas optoelectrónicos, *deep learning*, segmentación de la marcha.

Abstract

Cerebral palsy (CP) has been recognized as the most common motor disability in childhood (Bekteshi et al., 2023). Gait disorder in children suffering from this condition constitutes one of the main limitations to which they are subjected in their daily lives (Nowicki, 2020). This is why clinical gait analysis (CGA) has become a common practice in medical centers. It measures and analyzes multiple parameters, with the aim of identifying and understanding both different movement injuries and their causes (Armand et al., 2016).

As a complement to medical visual inspection, instrumented analysis has emerged, which is based on the use of electronic devices in order to provide a more complete and individualized study of each subject (Nowicki, 2020). Currently, there are many options of such devices offered on the market. Multi-camera machine vision is the gold standard technique in motion studies. (Tunca et al., 2017). Nevertheless, its high price and its need for large spaces and professional training, make it impossible to generalize instrumental gait analysis in the medical field (Gao et al., 2022).

For this reason, during this study a CGA was performed using inertial measurement units (IMU) and video systems. The aim is to demonstrate the potential of both techniques, which are much more economical than artificial vision, and which could make it possible to extend this type of practice to the hospital field. Thus, a database of 25 subjects was acquired, of which 17 were controls and 8 pediatric patients with CP. In addition, a deep learning model capable of automatically segmenting gait in its respective phases was developed.

The results obtained during the different training sessions showed that, in subjects without gait disorders, with 3-4 inertial sensors and a simple neural network architecture it is possible to perform high quality segmentation and analysis. Proof of this are the performance metrics obtained, showing an accuracy between 0.97 and 1. On the other hand, it was also possible to see how, in patients with cerebral palsy, it is necessary to train models that take into account all possible gait deviations, being impossible to use the same neural network in both groups of subjects. Finally, after all the tests carried out, the efficiency of both measurement techniques was demonstrated. Consequently, thanks to them, it is possible to achieve, in a more economical way, an objective and individualized study of each patient's gait.

Keywords

Cerebral palsy, instrumental gait analysis, inertial measurement units, optoelectronic systems, deep learning, gait segmentation.

Índice general

1. INTRODUCCIÓN	13
1.1 Contexto y motivación.....	13
1.2 Trabajos previos	14
1.3 Hipótesis y objetivos	15
1.4 Recursos utilizados	16
1.5 Organización de la memoria	16
2. PARÁLISIS CEREBRAL	18
2.1 Síntomas	18
2.2 Clasificación	18
3. PROCESO DE LA MARCHA	20
3.1 Marcha en sujetos sanos	20
3.2 Marcha en sujetos con parálisis cerebral	21
4. ANÁLISIS INSTRUMENTADO DE LA MARCHA	23
4.1 XSENS	23
4.2 OPENCAP	27
5. CONCEPTOS SOBRE DEEP LEARNING	30
5.1 Introducción.....	30
5.2 Conceptos básicos	30
5.3 Redes neuronales	32
5.4 Problema de desaparición del gradiente.....	35
5.5 Optimización y Regularización.....	36
5.6 Redes neuronales convolucionales	36
5.7 Redes neuronales recurrentes	38
6. MATERIALES Y MÉTODOS	40
6.1 Base de datos	40
6.2 Etiquetado de los datos.....	44
6.3 Pre-procesado de la base de datos.....	46
6.4 Red neuronal profunda para la segmentación de la marcha	48
7. RESULTADOS	52
7.1 Bloque 1. Entrenamientos con sujeto adulto sano.....	52
7.2 Bloque 2. Entrenamientos con sujeto adulto y pediátrico sano	57
7.3 Bloque 3. Entrenamiento con sujeto sano y patológico.....	60
7.4 Bloque 4. Entrenamiento con sujeto pediátrico con PC	62

8. DISCUSIÓN Y LIMITACIONES	66
8.1 Análisis de resultados en marcha de sujeto sano	66
8.2 Análisis de resultados en paciente con parálisis cerebral.....	70
8.3 Limitaciones.....	73
9. CONCLUSIONES Y LÍNEAS FUTURAS	74
9.1 Cumplimiento de los objetivos del trabajo de fin de grado.....	74
9.2 Conclusiones	74
9.3 Líneas futuras	75
BIBLIOGRAFÍA	76

Índice de figuras

Figura 2.1. Niveles del sistema GMFCS. Adaptación de (Jankovic et al., 2022).....	19
Figura 3.1. Fases del ciclo de la marcha. Adaptación de (Gao et al., 2022).....	21
Figura 4.1. Pantalla de configuración para la captura de movimiento en MVN Awinda 2023.0.	25
Figura 4.2. Representación de la distintas subfases de la fase de apoyo en el modelo biomecánico generado por MVN Awinda 2023.0. Adaptada de (<i>Xsens - MVN Gait report white paper.pdf</i> , s. f.).....	25
Figura 4.3. Representación de los parámetros espaciales. Adaptación de (<i>Gait Report - How to interpret data?</i> , s. f.).....	26
Figura 4.4. Representación gráfica en MotionCloud del ángulo de la rodilla, su media y su desviación estándar.....	27
Figura 4.5. Comparación entre el sistema Vicon y OpenCap (Uhlrich et al., 2022).....	28
Figura 4.6. Tablero de ajedrez usado para la calibración con OpenCap.....	28
Figura 4.7. Pose neutral correcta y posiciones incorrectas.....	29
Figura 5.1. Esquema de una neurona artificial.....	32
Figura 5.2. Arquitectura de una red neuronal profunda con múltiples capas (Parmar, 2018).	33
Figura 5.3. Funciones de activación. (a) Sigmoide. (b) Tangente hiperbólica. (<i>La Función de Activación</i> , s. f.).....	34
Figura 5.4. Función de activación ReLU (<i>La Función de Activación</i> , s. f.).....	34
Figura 5.5. Ejemplo de convolución sobre una entrada en 2D (Goodfellow et al., 2016).....	37
Figura 5.6. Efectos del padding en la dimensionalidad de salida.....	38
Figura 5.7. Estructura básica de una red neuronal recurrente.....	39
Figura 6.1. Disposición de las salas empleadas para la captura de datos. (a) Sala de rehabilitación del Hospital Universitario Río Hortega. (b) Sala “Leonardo Torres Quevedo”.	40
Figura 6.2. Adquisición de datos con sensores vestibulares. (a) Sensor Xsens MVN Awinda. (b) Disposición de los sensores en el cuerpo del sujeto mediante las cintas de velcro.....	41
Figura 6.3. Interfaz desarrollada en Python para el etiquetado de datos.....	44
Figura 6.4. Fases de la marcha (1-4) correspondientes al pie izquierdo.....	45
Figura 6.5. Representación de la dorsiflexión que se dio en sujetos normales y patológicos.	46
Figura 6.6. Arquitectura DNN desarrollada.....	49
Figura 6.7. Arquitectura CNN desarrollada.....	49

Figura 8.1. Matrices de confusión del Bloque 1 del clasificador derecho. Conjuntos de variables usados: (a) KA(RL), (b) KAH(R), (c) KAH(L).....	66
Figura 8.2. Matrices de confusión del Bloque 1 del clasificador izquierdo. Conjuntos de variables usados: (a) KA(RL), (b) KAH(L), (c) KAH(R).....	67
Figura 8.3. Sensores IMU necesarios para recolectar los conjuntos de variables (a) KA(RL), (b) KAH(R) y (c) KAH(L).....	68
Figura 8.5. Matrices de confusión del clasificador derecho del Bloque 4 utilizando como conjunto de variables: (a) KA(R) y (b) KA(L).....	71
Figura 8.6. Matrices de confusión del clasificador izquierdo del Bloque 4 utilizando etiquetas 1-6 y como conjunto de variables: (a) LOWERBODY(RL), (b) KAH(RL), (c) KA(RL) y (d) KA(L).....	72

Índice de tablas

Tabla 3.1. Patrones principales de marcha en sujetos con parálisis cerebral y sus respectivas características. Tabla simplificada y adaptada de (Papageorgiou et al., 2019).....	22
Tabla 4.1. Parámetros espaciales proporcionados en el informe de MotionCloud.	26
Tabla 6.1. Recomendaciones sobre la colocación de los dispositivos Xsens MVN Awinda.	41
Tabla 6.2. Descripción de la actividad a realizar durante la captura de datos.	43
Tabla 6.3. Características de los sujetos de control de la base de datos.	43
Tabla 6.4. Características de los sujetos pediátricos con PC de la base de datos.	43
Tabla 6.5. Archivos del lado izquierdo generados como paso previo al pre-procesado.	47
Tabla 6.6. Variables angulares proporcionadas por OpenCap.	48
Tabla 6.7. Distribución de las muestras en el Bloque 3 de entrenamientos.	50
Tabla 6.8. Subconjuntos de variables con información del miembro derecho e izquierdo.	51
Tabla 6.9. Subconjuntos de variables con información solo del miembro derecho.	51
Tabla 6.10. Subconjuntos de variables con información solo del miembro izquierdo.	51
Tabla 7.1. Entrenamientos del Bloque 1 del clasificador derecho.	52
Tabla 7.2. Subconjuntos creados para implementar <i>k-fold cross validation</i> para el Bloque 1.	53
Tabla 7.3. Entrenamientos <i>k-fold cross validation</i> para el clasificador derecho del Bloque 1.	54
Tabla 7.4. Comparación de entrenamientos <i>k-fold cross validation</i> para clasificador derecho en Bloque 1 según el lado usado como fuente de información.....	54
Tabla 7.5. Entrenamientos del Bloque 1 del clasificador izquierdo.	55
Tabla 7.6. Entrenamientos <i>k-fold cross validation</i> para el clasificador izquierdo del Bloque 1.	56
Tabla 7.7. Comparación de entrenamientos <i>k-fold cross validation</i> para clasificador izquierdo en Bloque 1 según el lado usado como fuente de información.....	56
Tabla 7.8. Entrenamientos del Bloque 2 del clasificador derecho.	57
Tabla 7.9. Comparación de entrenamientos para clasificador derecho en Bloque 2 según el lado usado como fuente de información.	58
Tabla 7.10. Entrenamientos del Bloque 2 del clasificador izquierdo.....	59
Tabla 7.11. Comparación de entrenamientos para clasificador izquierdo en Bloque 2 según el lado usado como fuente de información.	60
Tabla 7.12. Entrenamientos del Bloque 3 del clasificador derecho.....	60
Tabla 7.13. Comparación de entrenamientos para clasificador derecho en Bloque 3 según el lado usado como fuente de información.	61
Tabla 7.14. Entrenamientos del Bloque 3 del clasificador izquierdo.....	61

Tabla 7.15. Comparación de entrenamientos para clasificador izquierdo en Bloque 3 según el lado usado como fuente de información.	62
Tabla 7.16. Entrenamientos del Bloque 4 del clasificador derecho.....	62
Tabla 7.17. Subconjuntos creados para implementar <i>k-fold cross validation</i> para el Bloque 4.	63
Tabla 7.18. Comparación de entrenamientos k-fold cross validation para clasificador derecho en Bloque 4 según el lado usado como fuente de información.....	63
Tabla 7.19. Entrenamientos del Bloque 4 del clasificador izquierdo.....	64
Tabla 7.20. Distribución de las muestras en el Bloque 4 usando etiquetas 1-6 del clasificador izquierdo.	64
Tabla 7.21. Entrenamientos del Bloque 4 usando etiquetas 1-6 del clasificador izquierdo.	65

1. INTRODUCCIÓN

1.1 Contexto y motivación

El deterioro de la marcha, es una de las múltiples deficiencias motoras que condicionan la actividad diaria de los niños y niñas con parálisis cerebral (PC) (Nowicki, 2020). Aunque esta condición es permanente, las limitaciones que esta produce evolucionan y progresan con el paso del tiempo, necesitando de un monitoreo constante y ajustes en el tratamiento. Centrándonos en el problema de la marcha, en vez de buscar soluciones para aliviar y combatir los síntomas, resulta más eficaz determinar la causa subyacente, analizando los patrones de movimiento para, más adelante, reeducar a la misma, siendo este uno de los objetivos principales de los pacientes. Es por ello que, el análisis clínico de la marcha (*Clinical Gate Analysis*, CGA) se ha convertido en una práctica común dentro de los hospitales. Este mide y analiza múltiples parámetros (espacio-temporales, cinemáticos y cinéticos, entre otros) tratando de identificar y comprender a qué se deben las distintas desviaciones (Armand et al., 2016; Tunca et al., 2017). Además, ya no solo en el contexto de la parálisis cerebral, el CGA es indispensable para predecir y prevenir posibles caídas, ayudando así a disminuir la mortalidad y morbilidad asociadas (Tunca et al., 2017).

Antiguamente, los médicos se decantaban por una mera inspección visual, sin embargo, esto daba como resultado un análisis subjetivo y aislado incapaz de detectar correctamente los defectos, por ejemplo, en los movimientos de flexión/extensión, lo que hacía que muchos de los sujetos tuvieran que someterse a intervenciones quirúrgicas innecesarias o excesivas (Nowicki, 2020). Gracias a los avances técnicos, ha surgido una nueva modalidad de análisis, el análisis instrumentado, el cual se ayuda de dispositivos electrónicos para ofrecer un estudio mucho más completo y preciso del paciente. Así pues, a pesar de ser una técnica, a priori, más cara que la basada en la observación, este coste se compensa al verse reducido el número de intervenciones quirúrgicas y sesiones de rehabilitación necesarias (Wren et al., 2011).

En el mercado podemos encontrar una vasta selección de dispositivos para monitorizar la marcha, los cuales se pueden dividir entre directos e indirectos. Ninguno de ellos constituye el instrumento ideal, dado que lidian con ciertas limitaciones. Por ejemplo, dispositivos directos, generalmente basados en cámaras, tienden a filtrar los datos privados de los sujetos, no informan sobre ciertos parámetros de interés (como presiones) y no son idóneos para el análisis de actividades en el entorno natural. Por el contrario, los dispositivos indirectos son propensos a proporcionar datos menos precisos (Gao et al., 2022). Centrándonos en estos últimos, destacan las técnicas de distribución del estrés plantar, electromiografía y unidades de medición inercial (IMU), capaces de proporcionar información acerca de la fuerza de contacto, intenciones y movimiento del sujeto, respectivamente (Gao et al., 2022; Hebda-Boon et al., 2023). Una vez han sido registradas las señales, es común la extracción de parámetros de la marcha, mediante el uso de algoritmos especializados, necesarios para obtener una valoración más exhaustiva y cuidadosa del proceso (Gao et al., 2022). Existen distintos tipos de parámetros (espacio-temporales, mecánicos, eléctricos ...), tanto en el dominio temporal como frecuencial, cuyos valores difieren entre sujetos sanos y patológicos. Por ejemplo, las personas que sufren de PC tienden a tener una longitud de paso más corta y una velocidad de marcha más lenta en comparación con el resto de la población (Haakana et al., 2020).

Hoy en día, la visión artificial representa el *gold standard* en los sistemas de análisis de la marcha (Tunca et al., 2017). Esta tecnología se basa en el uso de cámaras de infrarrojo, las cuales son capaces de detectar y reconstruir la trayectoria seguida por una serie de marcadores localizados en puntos anatómicos concretos, siguiendo unas '*guidelines*' (Tunca

et al., 2017). Esta técnica es muy utilizada para capturar los movimientos de los actores en ciertas producciones cinematográficas, sin embargo, a pesar de su éxito y exactitud, su aplicación no se ha generalizado en el ámbito médico (Gao et al., 2022). La causa de esto puede ser su coste, lo que se le atribuye como problema principal, el cual es tan elevado que no todos los centros pueden permitírselo, así como su tamaño y la necesidad de formación profesional para su uso (Tunca et al., 2017). A modo de ejemplo cercano, en España menos de 10 centros médicos cuentan con sistemas de este tipo para la evaluación de pacientes, entre los que se encuentra el Hospital Universitario Virgen del Rocío (Sevilla) o el Hospital Infantil Universitario Niño Jesús (Madrid). Ninguno de estos centros está ubicado en la comunidad autónoma de Castilla y León.

En los últimos años ha crecido el interés por las unidades de medida inercial, dispositivos portátiles compuestos por acelerómetros, giroscopios y/o magnetómetros (Tunca et al., 2017). Demostración de ello es el aumento en el número de estudios existentes, cuya cifra se ha triplicado desde el 2009 (López-Nava & Muñoz-Meléndez, 2016). Estos, además de un precio más reducido, cuentan con otra serie de ventajas frente a los sistemas optoelectrónicos, los cuales son difícilmente portátiles, requieren de un elevado tiempo de colocación y limitan los datos recogidos a escasos pasos por ensayo. Concretamente, los dispositivos inerciales harían posible la adquisición continua de movimientos en lugares cotidianos para el paciente (el colegio, el hogar, durante otras actividades relevantes...), evitando así entornos artificiales, como los de un laboratorio, que lo condicionasen psicológicamente y pudiesen influir en su marcha (Bekteshi et al., 2023; Gao et al., 2022). A pesar de todas las virtudes nombradas anteriormente, las IMU no son aún un elemento cotidiano para el análisis de patrones de movimiento en el entorno clínico. Una de las razones la constituye la no disponibilidad de modelos de acceso abierto y suficientemente validados capaces de segmentar automáticamente la señal de marcha en sus respectivas fases.

Así pues, en este proyecto se va a abordar un doble análisis instrumentado de la marcha basado en la medida simultánea de los movimientos y fuerzas articulares, mucho más económico que los establecidos en el panorama médico actual. Concretamente, nos vamos a centrar en el estudio de la cinemática. Además, se elaborará un modelo de *deep learning* capaz de segmentar automáticamente los patrones de movimiento en sus respectivas fases de la marcha. En consecuencia, se obtendrá un conjunto de datos más completo, derivando en una visión más íntegra del proceso de la marcha, y, por lo tanto, en un mejor diagnóstico y tratamiento clínico para el paciente (Nowicki, 2020).

1.2 Trabajos previos

El reconocimiento de patrones de movimiento durante la marcha, así como la segmentación de la misma, constituyen un enfoque innovador que ha alcanzado un gran auge en estos últimos años. Gracias a los avances en el aprendizaje profundo, se ha conseguido superar y obtener unos mejores resultados que los obtenidos con las antiguas técnicas tradicionales, consistiendo estas en la segmentación de la silueta humana, seguida de la extracción y aprendizaje de características de la misma (Song et al., 2019). No obstante, la productividad obtenida era bastante baja, lo que obligó a la búsqueda de nuevas soluciones apoyándose en las últimas novedades del *deep learning* (Kidziński et al., 2019; Rifaat et al., 2022; Song et al., 2019).

Actualmente, existen dos enfoques diferentes para llevar a cabo esta clase de tareas. Por una parte, encontramos los modelos basados en el uso de parámetros espacio-temporales, y por otra, los métodos que se apoyan en el estudio de la forma y silueta humana extraída de una secuencia de imágenes (Song et al., 2019). Este último enfoque suele ser el empleado en las

labores de biometría, es decir, para la identificación de personas en misiones, por ejemplo, de videovigilancia (Rifaat et al., 2022). Es cierto que, este cometido, también puede realizarse mediante el reconocimiento de huellas dactilares o estructuras faciales. Sin embargo, la caracterización de la marcha supera limitaciones que afectan al rendimiento de estos anteriormente nombrados, como la larga distancia o condiciones de baja luminosidad (Song et al., 2019).

Nuestro proyecto en concreto no tiene que ver con una tarea de biometría, sino con un análisis individualizado de la marcha de cada paciente para poder proporcionar un diagnóstico, tratamiento, rehabilitación y atención más personalizada al mismo. Sin embargo, dado que en líneas generales los pasos necesarios para obtener tales fines son similares en ambos casos, es necesario comentar algunos trabajos de biometría con los que compararemos posteriormente el trabajo desarrollado en este trabajo de fin de grado (TFG). Así pues, en el 2019 Song et al. (2019) crearon Gait-Net, una arquitectura compuesta por dos redes convolucionales (*Convolutional Neural Network*, CNN), entrenadas de forma conjunta mediante imágenes de movimiento, para llevar a cabo labores de segmentación y clasificación de la marcha. Además, el año pasado, Rifaat et al. (2022) desarrollaron una nueva propuesta de estructura neuronal profunda, FCN-BiLSTM, basada en una red recurrente LSTM bidireccional, capaz de detectar dependencias temporales entre datos pasados y futuros, concatenada con una red totalmente convolucional. Este último trabajo guarda una mayor similitud con nuestro estudio actual, dado que utiliza como base de datos los parámetros espacio-temporales obtenidos a partir de, en este caso, un único sensor IMU, colocado en el centro del cinturón de los sujetos.

Centrándonos más en aplicaciones para la caracterización de sujetos patológicos, Borzì et al. (2023) hicieron, este mismo año, una red CNN para detectar eventos de congelación de la marcha en pacientes con la enfermedad de Parkinson. Cabe destacar que, al igual que Rifaat et al. (2022), emplearon un único sensor inercial ubicado en la cintura.

Ahora bien, enfocándonos en la aplicación y objetivo de este TFG, Martindale et al. (2021) propusieron una red capaz de predecir la fase de cada pie, así como distinguir entre ocho posibles actividades. Para ello, combinaron una red convolucional, para detectar los bordes, junto con redes GRU, una modalidad de red recurrente, la cual fue entrenada con los datos recogidos por cuatro sensores inerciales. Por último, un estudio mucho más similar al nuestro sería el de Kreuzer & Munz (2021), el cual hace uso de redes CNN y LSTM para clasificar entre cinco fases de la marcha, a partir de los datos recopilados por once dispositivos IMU.

1.3 Hipótesis y objetivos

En este TFG se plantea la hipótesis de si es posible realizar un análisis de la marcha a partir de dispositivos de bajo coste y herramientas de software libre con la suficiente relevancia clínica como para producir avances en este ámbito, capaces de ayudar a la generalización y expansión de análisis instrumentados en ambientes hospitalarios.

El objetivo principal de este proyecto es lograr desarrollar un sistema de aprendizaje automático capaz de segmentar la marcha en sus distintas fases y realizar pruebas de funcionamiento con sujeto sano y con paciente pediátrico con parálisis cerebral. Con el fin de alcanzar este objetivo, se ha dividido este Trabajo de Fin de Grado en pequeños objetivos específicos los cuales, según se vayan cumpliendo con éxito, nos acercarán al logro del principal. Estos son los siguientes:

- i. Realizar una profunda búsqueda bibliográfica hasta conseguir una completa documentación acerca del análisis instrumentado de la marcha en el panorama actual de la parálisis cerebral.
- ii. Crear una base de datos multimodales (vídeo e IMUs) de patrones de marcha de sujetos sanos y pacientes con parálisis cerebral.
- iii. Crear una herramienta en Python para facilitar el etiquetado manual de las fases de la marcha.
- iv. Realizar pruebas de clasificación de las fases de la marcha con redes neuronales de aprendizaje profundo.
- v. Analizar y comparar los resultados alcanzados con los obtenidos en estudios similares ya publicados.
- vi. Extraer conclusiones a partir de los datos obtenidos durante la realización de las diferentes pruebas, limitaciones y posibles mejoras futuras.

1.4 Recursos utilizados

Durante el desarrollo de este trabajo se hizo uso de los siguientes recursos de hardware y software:

Hardware

- Xsens MVN Awinda Starter Pack®: formado por 17 sensores inerciales inalámbricos junto con las correspondientes cintas ajustables necesarias para colocarlos sobre el cuerpo del sujeto. Estos tienen una frecuencia de muestro máxima de 60 Hz y un rango máximo de 25 metros.
- 2 Apple iPad®.
- Ordenador con un procesador i7, 8GB de memoria y una tarjeta GTX 2080.

Software

- Xsens Software MVN Awinda 2023.0.
- Python 3.10.11.
- Aplicación OpenCap versión 1.6.

1.5 Organización de la memoria

Este documento está dividido en 9 capítulos. Durante este primer capítulo introductorio se han presentado los conceptos básicos sobre los que se sustenta el presente TFG junto con el contexto actual y el problema en cuestión, así como la hipótesis de trabajo, los objetivos y los recursos empleados. Se completa con los restantes 8 capítulos explicados a continuación:

En el **Capítulo 2 “Parálisis Cerebral”** se presenta este problema de salud, haciendo hincapié en su sintomatología y posibles formas de clasificar a los pacientes que sufren de esta condición.

En el **Capítulo 3 “Proceso de Marcha”** se explica de forma detallada el movimiento de marcha, dividiéndolo en sus respectivas fases y subfases. Además, se exponen las desviaciones de la marcha más comunes entre los sujetos con PC.

En el **Capítulo 4 “Análisis Instrumentado de la Marcha”** se presenta el panorama actual en el que se mueve este tipo de análisis, explicando y comparando los sistemas utilizados durante el desarrollo de este TFG con el método *gold standard* a día de hoy.

En el **Capítulo 5 “Conceptos sobre Deep Learning”** se explican de forma detallada los conceptos básicos de este campo, insistiendo en los dos tipos de redes más utilizadas en la literatura existente para lograr el objetivo principal del proyecto.

En el **Capítulo 6 “Materiales y Métodos”** se muestra la base de datos recopilada para la ejecución de este TFG, junto con el pre-procesamiento de la misma. Asimismo, se describen las redes desarrolladas y los entrenamientos aplicados para la clasificación de movimientos de la marcha.

En el **Capítulo 7 “Resultados”** se comentan los resultados obtenidos durante las distintas pruebas realizadas, para cada uno de los diferentes clasificadores entrenados.

En el **Capítulo 8 “Discusión y Limitaciones”** se analizan y comparan los resultados obtenidos en este estudio con otros existentes similares. A su vez, se comentan los hallazgos encontrados en este TFG, así como las limitaciones e inconvenientes encontrados durante la ejecución de este proyecto.

En el **Capítulo 9 “Conclusiones y Líneas Futuras”** se presentan las conclusiones extraídas en este trabajo y se muestran las posibles líneas futuras a partir de las cuales se podría dar continuidad a este estudio.

2. PARÁLISIS CEREBRAL

Como muchos especialistas en el tema apuntan, la parálisis cerebral no es una enfermedad aislada como tal, sino más bien un conjunto de desórdenes que dan como resultado final una malformación cerebral (Nowicki, 2020). Hasta 1980 se consideraba la asfixia durante el nacimiento como la única causa de PC. Sin embargo, gracias a la investigación biomédica se confirmó que esta solo era una de las muchas etiologías posibles (Tugui & Antonescu, 2013). Así pues, actualmente se sabe que también puede desarrollarse en la infancia debido a afecciones como traumatismos craneales, infecciones cerebrales o, incluso ictericias no tratadas. En consecuencia, la gran variedad de fenotipos con sintomatologías comunes pero idiopáticas diferentes vuelven a la PC un concepto difícil de limitar. A pesar de ello, algunos expertos se han atrevido a desarrollar una definición (Richards & Malouin, 2013):

“La PC es un trastorno del desarrollo del movimiento y la postura, que causa limitaciones en la actividad como consecuencia de alteraciones no progresivas del cerebro fetal o infantil, las cuales también pueden afectar la sensibilidad, la percepción, la cognición, la comunicación y el comportamiento.”

Dada su prevalencia mundial de 1,6 a 3,4 por cada 1000 nacidos vivos, esta patología es considerada la discapacidad motora más frecuente de la infancia (Bekteshi et al., 2023). De ahí la importancia de la investigación y trabajo en cuanto al diagnóstico, tratamiento y monitorización.

2.1 Síntomas

Como resultado de la lesión cerebral, más o menos extensa y grave, los sujetos pueden tener que lidiar en su día a día con alguno de los siguientes síntomas (Prieto, 2016):

- **Trastornos motores.** Son múltiples las formas en las que estos se manifiestan, sin embargo, todos ellos condicionan en mayor o menor medida la actividad diaria de aquellas personas que presentan PC. En el siguiente apartado, serán presentadas en mayor profundidad todo ese abanico de posibles lesiones.
- **Trastornos sensoriales.** Estos suelen afectar al campo visual y auditivo.
- **Alteraciones del lenguaje.** Dadas las lesiones en áreas cerebrales concretas, así como los déficits motores de la musculatura encargada del habla, estos niños suelen sufrir complicaciones en cuanto a su inserción y comunicación en la sociedad.
- **Episodios epilépticos.** Surgen como consecuencia de la afección neurológica.

Como consecuencia de todo ello, ya presenten solo una o una mezcla de las manifestaciones arriba nombradas, estas personas pueden presentar graves retrasos intelectuales, así como dificultades para alcanzar cierto grado de autonomía personal en tareas tan sencillas como pueden ser comer, vestirse o atarse los zapatos (Prieto, 2016). De todas ellas este TFG se va a centrar en el déficit motor, más concretamente, en los obstáculos que este supone para la marcha en esta clase de pacientes.

2.2 Clasificación

Clínicamente, los pacientes con PC pueden ser clasificados de tres formas, según la anatomía de la lesión, la deficiencia motora y las capacidades funcionales (Armand et al., 2016; Tugui & Antonescu, 2013).

Centrándonos en el primer tipo de clasificación, anatómicamente, encontramos los siguientes tipos de déficits motores (Tugui & Antonescu, 2013):

- **Hemiplejia.** Pérdida de control de la mitad del cuerpo.
- **Paraplejia.** Pérdida de control de las extremidades inferiores.
- **Tetraplejia.** Pérdida de control tanto de las extremidades inferiores como superiores.
- **Triplejia.** Pérdida de control de las dos extremidades inferiores y una superior.
- **Monoplejia.** Pérdida de control de una sola extremidad.

La paraplejia constituye la forma más común, seguida de la hemiplejía y tetraplejía, las cuales ocupan el segundo y tercer puesto, respectivamente (Tugui & Antonescu, 2013).

Los trastornos motores pueden ser espásticos, atetósicos, distónicos, hipotónicos y atáxicos, entre otros. De todos ellos, la espasticidad es el más frecuente, estando presente en más del 60% de los casos de parálisis cerebral (Armand et al., 2016). Este es consecuencia de una alteración en el tracto piramidal, es decir, de la vía motora encargada de los movimientos voluntarios. El resultado consiste en un aumento del tono muscular junto con fatiga, así como la falta de coordinación y destreza y la aparición de contracturas (Armand et al., 2016).

Por último, atendiendo a la tercera forma, el Sistema de la Clasificación de la Función Motora Gruesa (*Gross Motor Function Classification System, GMFCS*) distingue cinco niveles de gravedad en función de las limitaciones funcionales, la necesidad de uso de dispositivos auxiliares de la marcha (muletas, bastones, andadores) o sillas de ruedas (manuales o eléctricas, autopropulsadas o no) (Jankovic et al., 2022). Estos son:

- **Nivel I.** Anda sin limitaciones.
- **Nivel II.** Anda con limitaciones.
- **Nivel III.** Anda con ayuda de un dispositivo de movilidad con sujeción manual.
- **Nivel IV.** Autonomía para la movilidad con limitaciones; puede usar sistemas de propulsión a motor.
- **Nivel V.** Transportado en una silla de ruedas manual.



Figura 2.1. Niveles del sistema GMFCS. Adaptación de (Jankovic et al., 2022).

3. PROCESO DE LA MARCHA

3.1 Marcha en sujetos sanos

Se conoce como ciclo de marcha o zancada al conjunto de eventos que tienen lugar entre dos repeticiones sucesivas de uno cualquiera de los sucesos de la marcha (Kharb et al., 2011; Tunca et al., 2017). Cada ciclo puede dividirse en dos fases, la de apoyo y la de balanceo, en función de si el pie de referencia, generalmente el derecho, soporta el peso o no. La primera de ellas comprende aproximadamente el 62% del ciclo, extendiéndose desde el contacto inicial del pie con el suelo (IC) hasta el momento en el que lo abandona (FO), es decir, cuando hay un ascenso completo de la carga. En cambio, en la fase de balanceo, este pie permanece en el aire durante el 38% del tiempo restante (Tunca et al., 2017). Los eventos IC y FO se suelen nombrar como talón y dedo del pie, respectivamente. No obstante, estos términos alternativos solo tienen un significado firme en marcha no patológica, dado que, en sujetos enfermos podrían no ser esas partes del pie las primeras en ser apoyadas y/o últimas en ser levantadas (Gao et al., 2022; Tunca et al., 2017).

A su vez, cada una de estas fases puede subdividirse en varias etapas, tal y como se muestra en la Figura 3.1. Para la fase de apoyo distinguimos los siguientes eventos (Gao et al., 2022):

- **Contacto inicial (IC)** (0-3%). Como se comentó en el anterior párrafo, en un patrón de marcha normal, en esta fase el talón es el que establece contacto con el suelo. Así pues, para que esto se ejecute de forma correcta, se genera una flexión dorsal del tobillo, pequeña o nula extensión de la rodilla y flexión de la cadera.
- **Respuesta a la carga** (3-12%). El pie de referencia está completamente apoyado en el suelo y, el otro restante, al borde de levantarse del suelo. Además, la rodilla se encuentra flexionada para mantener el equilibrio y amortiguar el impacto, produciendo una desaceleración.
- **Fase media de apoyo** (12-31%). Esta etapa comienza con el abandono del suelo del miembro contralateral (el cual comienza su fase de balanceo), de forma que ahora toda la carga recae sobre la pierna de referencia. Para hacer esto posible, la cadera y la rodilla deben estabilizarse.
- **Fase final de apoyo** (31-50%). El talón de referencia se despega del suelo, de forma que el centro de presión plantar queda desplazado hacia la parte anterior del pie. Es decir, el cuerpo sobrepasa la línea vertical, generando una aceleración.
- **Foot Off (FO)** (50-62%). También se la suele denominar como etapa Pre-Swing. El centro de presión sigue desplazándose hacia delante hasta que, en marcha no patológica, los dedos se levantan del suelo.

A su vez, la fase de balanceo puede dividirse en 3 etapas (Gao et al., 2022):

- **Fase inicial de balanceo** (0-13%). El objetivo de esta etapa es adelantar la pierna, a la vez que se consigue una separación pie-suelo suficiente. Para ello, es necesario el trabajo de la musculatura pretibial y de la cadera.
- **Fase media de balanceo** (13-25%). Durante la misma, el pie de referencia alcanza la mayor distancia de separación con el suelo, gracias a la contracción de los flexores dorsales.
- **Fase final de balanceo** (25-38%). El movimiento de la pierna se desacelera para posicionarla correctamente antes de comenzar la etapa de IC. Para este fin, la rodilla debe estar completamente extendida y el pie tener una posición neutra con respecto a la pierna.

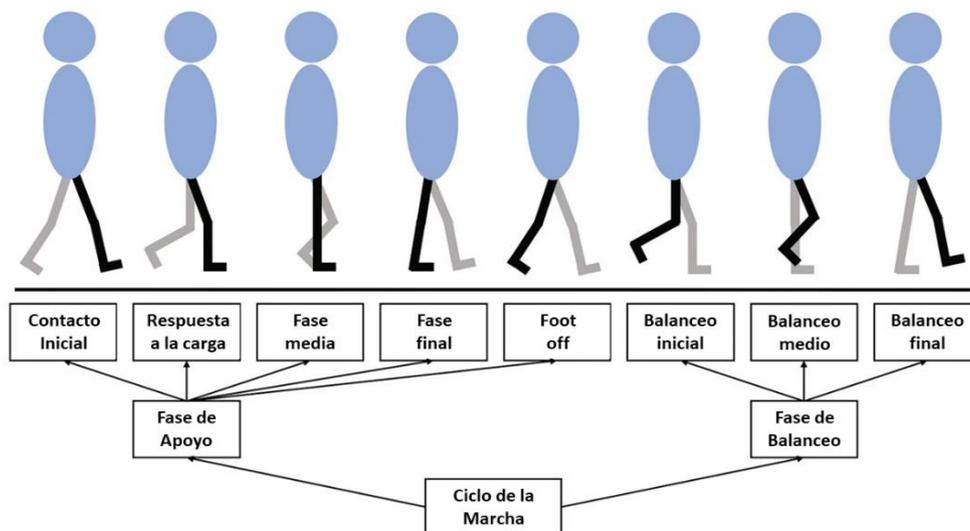


Figura 3.1. Fases del ciclo de la marcha. Adaptación de (Gao et al., 2022).

3.2 Marcha en sujetos con parálisis cerebral

Perry y Burnfield especificaron algunas condiciones necesarias para que la marcha fuese efectiva (Perry et al., 1992). Estas eran: estabilidad durante la fase de apoyo, libertad del pie en el balanceo, apropiada distancia entre pasos y adecuada colocación del pie para el contacto inicial. Las personas con parálisis cerebral tienen lesionado parte de su sistema nervioso, lo que deriva en el incorrecto funcionamiento del sistema locomotor y, en consecuencia, en dificultades para cumplir y mantener dichos requisitos al caminar.

Es por ello por lo que, las fases y etapas anteriormente descritas pueden ejecutarse incorrectamente, o simplemente no darse, en pacientes con PC. De entre todas las posibles desviaciones de la marcha, la hemiplejía y paraplejía espástica constituyen los trastornos más frecuentes (Armand et al., 2016). En 2019, se realizó una revisión sistemática acerca de las distintas clasificaciones de la marcha en pacientes pediátricos con parálisis cerebral. Se llegó a la conclusión de la existencia de 6 patrones a los que denominaron: "genu recurvatum", "pie caído", "equino verdadero", "marcha de salto", "equino aparente" y "marcha agachada" (Papageorgiou et al., 2019). Las características más notables de los mismos vienen descritas en la Tabla 3.1:

PATRÓN		CARACTERÍSTICAS
Genu Recurvatum	Tobillo	La extensión excesiva de la rodilla afecta negativamente al control motor del tobillo, pudiendo dar flexión plantar o dorsal.
	Rodilla	Hiperextensión (extensión > 180°) en la fase de apoyo.
Pie Caído	Tobillo	Dificultad para realizar la flexión dorsal. El IC suele darse con la punta del pie.
	Rodilla	Para evitar el arrastre del pie durante la fase de balanceo, los pacientes suelen levantar el muslo mientras flexionan la rodilla.

Equino Verdadero	Tobillo	Flexión plantar durante la fase de apoyo, es decir, se apoya únicamente la parte anterior del pie (marcha en puntillas).
	Rodilla	Se extiende totalmente, pudiendo llegar a la hiperextensión.
	Pelvis	Normal o ligeramente inclinada hacia delante.
Marcha de Salto	Tobillo	Tobillo en equino, particularmente en la fase final de apoyo.
	Rodilla y Cadera	Flexión excesiva de rodilla y cadera en la fase inicial de apoyo, seguida de la extensión de las mismas en la fase final.
	Pelvis	Pelvis normal o ligeramente anteriorizada.
Equino Aparente	Tobillo	Movimiento dentro del rango normal, aunque el talón no toca el suelo.
	Rodilla	Hiperflexión durante la fase de apoyo.
	Pelvis	Normal o ligeramente inclinada hacia delante.
Marcha Agachada	Tobillo	Flexión dorsal excesiva durante la fase de apoyo.
	Rodilla y Cadera	Hiperflexión de ambas.
	Pelvis	Normal o con retroversión.

Tabla 3.1. Patrones principales de marcha en sujetos con parálisis cerebral y sus respectivas características. Tabla simplificada y adaptada de (Papageorgiou et al., 2019).

Al observar la tabla, es posible afirmar que las diferencias entre los distintos patrones son muy leves, ya que muchos de ellos son una combinación de otros, donde pueden distinguirse tanto elementos comunes como características más propias. He ahí la razón y necesidad de añadir un análisis instrumentado adicional al trabajo visual del médico o rehabilitador, ya que siempre va a existir cierto grado de subjetividad en cuestiones tan mínimas y precisas como es un patrón de movimiento.

4. ANÁLISIS INSTRUMENTADO DE LA MARCHA

Tal y como se comentó en el Capítulo 3. Proceso de la marcha y en el Capítulo 1. Introducción, el análisis instrumentado de la marcha (*Instrumental Gait Analysis*, IGA) constituye un importante y necesario complemento de la observación visual médica. Por ejemplo, el IGA adquiere una máxima consideración en casos en los que hay que distinguir entre un equino aparente y uno verdadero, donde las diferencias son mínimas como para ser diferidas a simple vista (Kumar et al., 2021). Luego, el uso conjunto de ambos permite identificar en mejor medida patrones complejos del movimiento, derivando en una consulta y tratamiento mucho más individualizado y completo para el paciente.

Desde principios de los 80, década en la que se instauró este nuevo concepto, médicos e ingenieros han colaborado y fusionado sus esfuerzos para obtener sistemas cada vez más precisos, manejables y rentables (States et al., 2021). A día de hoy, destaca el sistema de captura de movimiento Vicon (Oxford Metrics Limited, England), el cual es considerado el líder de entre todos los posibles métodos de análisis y recopilación de datos biomecánicos. Esto lo consigue mediante el uso de un conjunto de cámaras de infrarrojo capaces de localizar y hacer un seguimiento de una serie de marcadores reflectantes colocados en puntos articulares y segmentos corporales del sujeto. Su elevado coste es considerado su principal desventaja y, es por ello y por su complejidad de uso, por lo que su utilización no está tan extendida en la clínica médica.

Así pues, en este trabajo de fin de grado se va a emprender un estudio de la marcha a partir del empleo de unidades de medición inercial y sistemas de captura vídeo, ambas modalidades mucho más asequibles que Vicon.

4.1 XSENS

Xsens (Movella, USA) es un sistema de análisis de movimiento basado en unidades inerciales, es decir, dispositivos portátiles compuestos por acelerómetros, giroscopios y/o magnetómetros. Así pues, es capaz de proporcionar un estudio completo y objetivo acerca del movimiento del paciente, en este caso particular, sobre su marcha. Concretamente, facilita parámetros de posición, velocidad y aceleración de los segmentos corporales, así como información tridimensional sobre cada una de las articulaciones y sus respectivos centros de masa.

Para este TFG se empleará el sistema de hardware Xsens MVN Awinda. Más adelante, en el Capítulo 6. Materiales y Métodos se explicará de forma detallada el protocolo de colocación de estos sensores y cómo los datos proporcionados por estos han sido utilizados.

4.1.1 Comparación con Vicon

Los sistemas optoelectrónicos, como lo es Vicon, requieren de espacios calibrados y concretos que reducen su portabilidad, lo cual hace que su uso no se haya extendido tanto como cabría esperar, teniendo en cuenta que proporcionan información muy precisa y completa (Blair et al., 2018; Gao et al., 2022).

Así pues, los sensores inerciales Xsens vienen acompañados de una serie de ventajas con respecto a este actual sistema *gold standard*, lo que les convierte en un importante competidor de mercado. Más concretamente, Xsens está libre de cualquier tipo de restricción espacial y condiciones de iluminación, pudiendo tomar los datos fuera del ambiente de laboratorio. Adicionalmente, hay que considerar que su precio no supera los

5000 €, lo cual contrasta enormemente con el elevado coste que caracteriza a Vicon. (*Xsens-MVN Gait report white paper.pdf*, s. f.).

De este modo, los sensores inerciales superan muchas de las limitaciones de los equipos optoelectrónicos, al mismo tiempo que, tal y como muchos estudios permiten confirmar, obtienen un análisis cinemático muy similar e igual de válido que el proporcionado por estos últimos (Blair et al., 2018).

4.1.2 Proceso de adquisición de datos

Adquirir datos de movimiento con los dispositivos inerciales Xsens es un proceso cómodo y fácil de seguir. Sin embargo, previamente es necesario tener en cuenta una serie de consideraciones acerca de la colocación de los sensores ya que, en caso contrario, los datos recogidos podrían no representar correctamente la realidad. Además, es importante mencionar que para el estudio de la marcha bastaría con el uso de 7 de los 17 sensores MVN Awinda disponibles, siendo estos los ubicados en la zona de la pelvis (1), parte superior (2) e inferior de la pierna (2) y pie (2) (*Xsens - MVN Gait report white paper.pdf*, s. f.). Sin embargo, para este proyecto se optó por hacer uso de la totalidad de IMUs durante las capturas de datos, dado que el movimiento de los miembros superiores del cuerpo podría proporcionar información valiosa capaz de ayudar a la red neuronal en su tarea de clasificación.

El primer paso consiste en la colocación de los sensores. Hay que mencionar que las unidades de medición inercial se caracterizan por medir directamente orientaciones y no posiciones, gracias a lo cual, la precisión en el posicionamiento de los sensores no tiene mucha importancia, permitiendo así eliminar el sesgo introducido por el operador durante las diferentes sesiones. No obstante, para evitar el movimiento de los mismos durante la ejecución del movimiento, deben fijarse fuertemente con las cintas de velcro. Asimismo, se recomienda posicionarlos en las zonas planas del cuerpo y con menor cantidad de tejido graso, idealmente apoyados sobre el hueso (*Xsens - MVN Gait report white paper.pdf*, s. f.).

Tras el posicionamiento, se inicia una captura de movimiento, configurada a partir de los datos de altura y longitud de pie del sujeto, tal y como se muestra en la Figura 4.1. Además, en la foto puede verse cómo se indica si los sensores están correctamente encendidos, marcándolos con el color verde.

Por último, como paso previo a la captura de datos, es necesario calibrar las IMU. Existen varias opciones de calibración, sin embargo, para este proyecto se optó por la recomendada por los propios sensores, *Npose + walk*. Esta se basa en lo siguiente:

- El sujeto comienza realizando la pose neutra (*Npose*) durante 3 segundos, erguido con la mirada al frente, separando los pies ligeramente y con los brazos apoyados a ambos lados del cuerpo.
- Pasados esos segundos, el programa pedirá al sujeto que camine por la sala durante otros 10 segundos, a un ritmo cómodo y balanceando los brazos de manera natural.
- Por último, se lo solicitará al sujeto que regrese al punto inicial y que vuelva a sostener la posición neutra hasta que finalice la calibración.

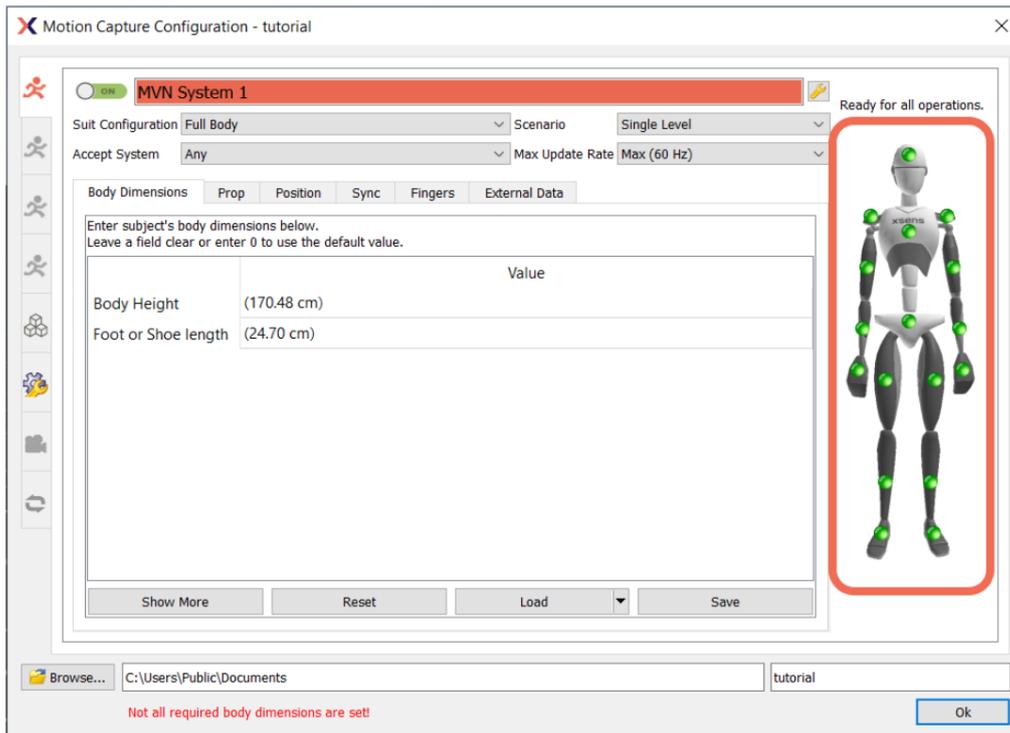


Figura 4.1. Pantalla de configuración para la captura de movimiento en MVN Awinda 2023.0.

Una vez hechos todos estos pasos ya es posible la grabación de movimientos. Estos son mostrados por la pantalla en tiempo real, a partir del uso de un modelo biomecánico, mostrado en la Figura 4.2. De esta forma, es posible detectar errores de colocación o calibrado de los sensores, en el caso de que el movimiento del modelo difiera del realizado por el participante en la realidad. Una vez completada la captura, esta puede ser guardada y así ser visualizada siempre que se desee.

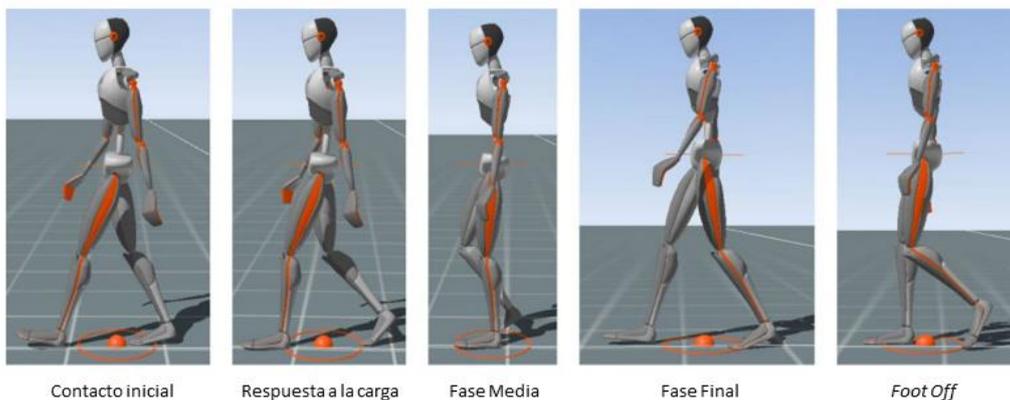


Figura 4.2. Representación de las distintas subfases de la fase de apoyo en el modelo biomecánico generado por MVN Awinda 2023.0. Adaptada de (Xsens - MVN Gait report white paper.pdf, s. f.).

4.1.3 Análisis de datos en la nube

Los archivos *‘.mvn’* generados durante las capturas con los sensores Xsens MVN Awinda permiten ser posteriormente analizados en su propia nube, MotionCloud, obteniendo así un informe de la marcha para cada uno de los sujetos. Es cierto que, para que esto sea posible,

cada individuo debe dar, como mínimo, 10 pasos sucesivos en línea recta (*Xsens - MVN Gait report white paper.pdf*, s. f.).

Una vez subido el archivo a la nube, el informe generado proporciona dos tipos de datos:

- **Parámetros de la marcha.** La información proporcionada viene resumida en la siguiente tabla.

Tipos de parámetros	Descripción
Generales	Velocidad (m/s), cadencia (pasos/min), número de pasos, duración (s), y distancia (m).
Contador de eventos de contacto	Hace referencia al número de contactos registrados, es decir, aquellos momentos en los que el pie, tras estar levantado, se apoya de nuevo en el suelo. Además, también registra el número de ocasiones en las que el miembro inferior despegó para comenzar la fase de swing, dejando de estar en contacto con el suelo. Para todo ello distingue entre pie izquierdo y pie derecho, así como si el contacto/despegue se realizó con el talón o con la punta.
Espaciales	Longitud de paso, ancho de paso y longitud de zancada.
Temporales	Da información acerca de la duración de: ciclo de la marcha, paso, fase de balanceo, fase de apoyo y cada una de las distintas subfases.

Tabla 4.1. Parámetros espaciales proporcionados en el informe de MotionCloud.

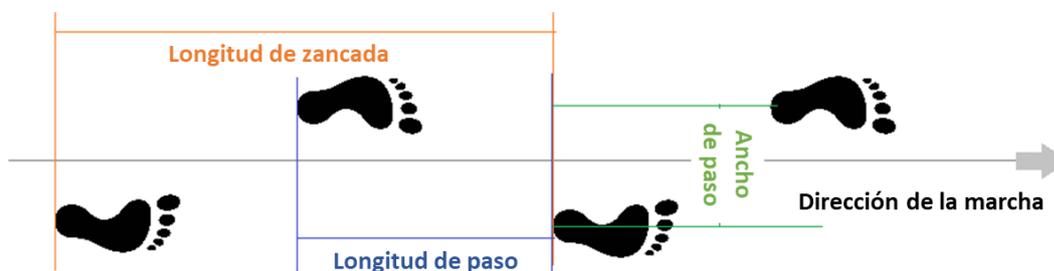


Figura 4.3. Representación de los parámetros espaciales. Adaptación de (*Gait Report - How to interpret data?*, s. f.).

- **Gráficos de la marcha.** Esta segunda sección del informe elabora gráficas acerca de cómo varían tanto los ángulos de las distintas articulaciones del miembro inferior, como la orientación de la pelvis, el centro de masa y la aceleración de la pierna a lo largo de los sucesivos ciclos de la marcha. Se muestran los parámetros para todos los ciclos de la marcha registrados, distinguiéndose entre pie izquierdo (azul) y pie derecho (naranja), pudiéndose, además, visualizar la media y la desviación estándar de los mismos. A mayores, en la parte inferior, se presenta una tabla que recoge los valores máximo y mínimo registrados tanto en la fase de apoyo como en la de balanceo, para el respectivo parámetro seleccionado.

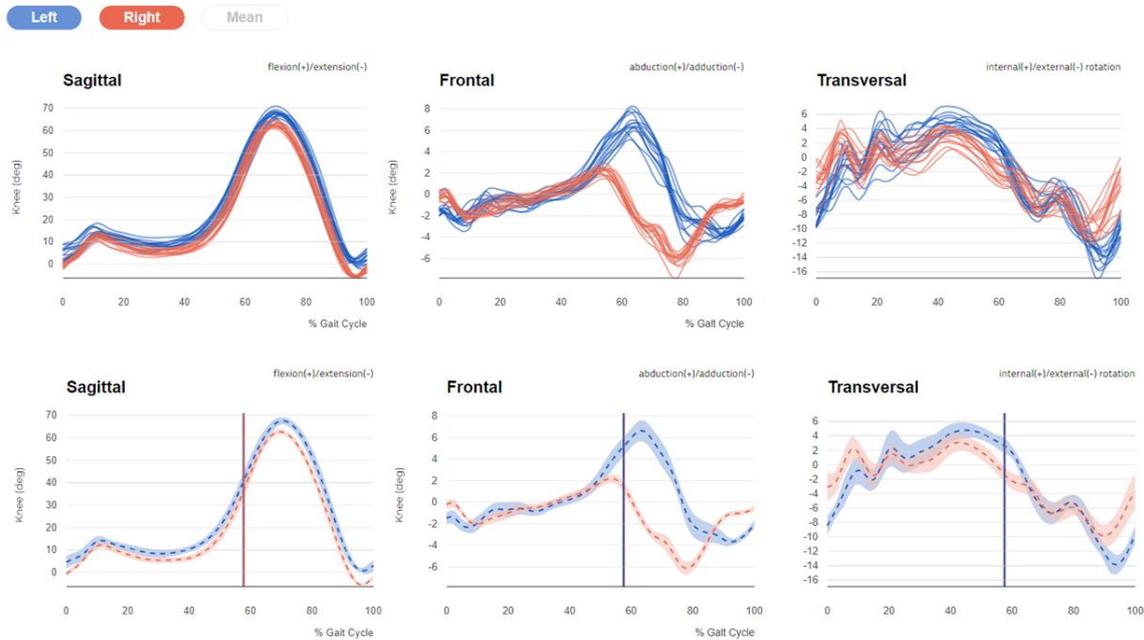


Figura 4.4. Representación gráfica en MotionCloud del ángulo de la rodilla, su media y su desviación estándar.

4.2 OPENCAP

OpenCap es una plataforma de código abierto, desarrollada por miembros del Laboratorio de Biomecánica Neuromuscular de la Universidad de Stanford, capaz de realizar un análisis del movimiento no sólo cinemático, sino también dinámico, todo ello a partir de videos capturados desde dispositivos tipo *Smartphone* o *tablet*. Las ventajas de este método recaen en que únicamente requiere de 2 dispositivos iOS para realizar su función, lo que lo convierte en un sistema económico y portátil. Además, es rápido y fácil de utilizar, dado que los videos se procesan automáticamente en la nube, sin necesidad de algún tipo de experiencia, *hardware* o *software* especializados (Uhlrich et al., 2022).

4.2.1 Comparación con Vicon

Al comienzo del capítulo se habló del sistema de análisis Vicon, basado en la visión artificial, cámaras de infrarrojo y marcadores anatómicos. Tal y como es sabido, el estudio cuantitativo de la dinámica, no solo de la marcha, sino también de otros muchos movimientos, es necesario para alcanzar un diagnóstico más profundo y un tratamiento más personalizado (Hebda-Boon et al., 2023). Sin embargo, equipos como el anterior dificultan que este proceso se extienda a la práctica médica cotidiana. Ello es debido a su necesidad de laboratorios de grandes dimensiones y una formación especializada, así como a su elevado coste, donde el conjunto de hardware y software puede llegar a sobrepasar los 30.000 € (Uhlrich et al., 2022).

Si comparamos los sistemas Vicon con la novedosa aproximación seguida por Stanford, podemos ver los múltiples beneficios que este último ofrece. No obstante, se necesitarán estudios adicionales para valorar si con OpenCap se consiguen unos resultados tan válidos y útiles como los proporcionados por los primeros. Tal y como puede verse en la Figura 4.5, OpenCap permite reducir enormemente el tiempo de adquisición y procesado de datos, ya que no necesita de una colocación previa de marcadores en el cuerpo del paciente y, además, la calibración no sobrepasa los 5 minutos. Asimismo, puesto que la información recopilada

se procesa automáticamente en la nube, no requiere de ninguna formación previa ni de ninguna clase de software concreto. Adicionalmente, a todas estas ventajas anteriores, hay que sumarle que únicamente requiere de 2 dispositivos móviles para la captura de datos, lo que hace que su precio no ascienda de los 700 € (Uhlrich et al., 2022).

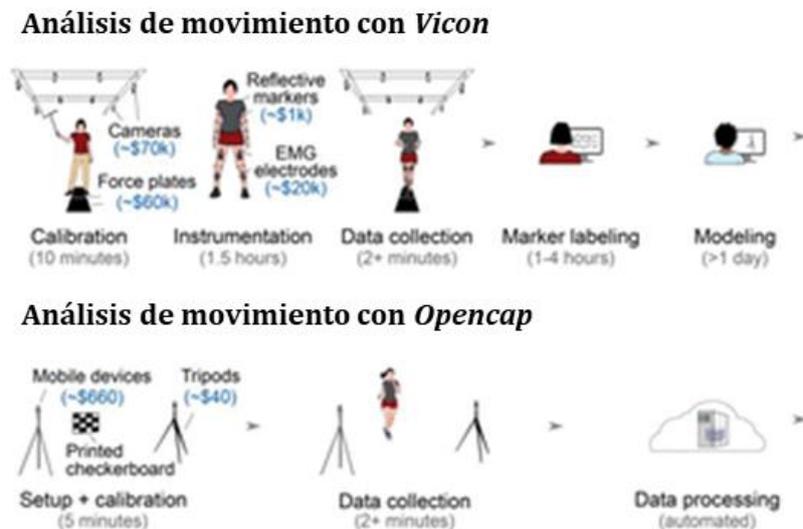


Figura 4.5. Comparación entre el sistema Vicon y OpenCap (Uhlrich et al., 2022).

4.2.2 Proceso de adquisición de datos

Durante este apartado se explicará brevemente cómo es el procedimiento seguido para capturar datos a partir de la plataforma OpenCap. Previamente es necesario haber ubicado los *smartphones* correctamente, comprobando que ningún segmento corporal se escape del área de visión de los dispositivos durante los movimientos. Para evitar este último problema, se recomienda evitar los planos sagitales y posicionar las cámaras formando un ángulo de 30-45°. Una vez verificado este punto, ya puede dar comienzo la adquisición.

Primero, es necesario un proceso de calibración del sistema. Para ello, se hace uso de un tablero de ajedrez, el cual debe posicionarse centrado, perpendicular al suelo y ser visible desde todas las cámaras. Además, para obtener unos resultados óptimos, el tablero no debe situarse a más de cinco metros de los *smartphones*.

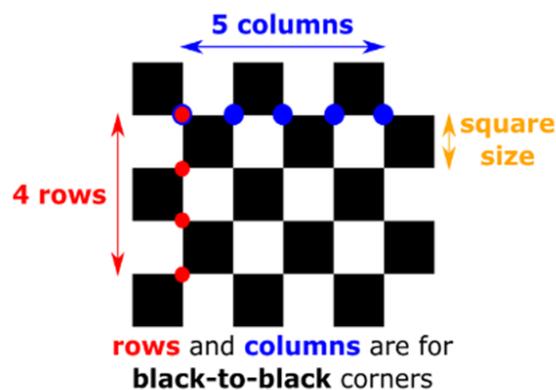


Figura 4.6. Tablero de ajedrez usado para la calibración con OpenCap.

Tras este paso, la aplicación web de OpenCap muestra una pantalla para completar datos, como la altura (m) y peso (Kg) del sujeto, imprescindibles para comenzar con la grabación del mismo. La primera posición capturada es la pose neutral, necesaria para calibrar a la persona antes de comenzar con los movimientos. Se recomienda que el sujeto lleve ropa entallada y que contraste con la sala, con el fin de que las cámaras lo visualicen claramente y sin ningún inconveniente.

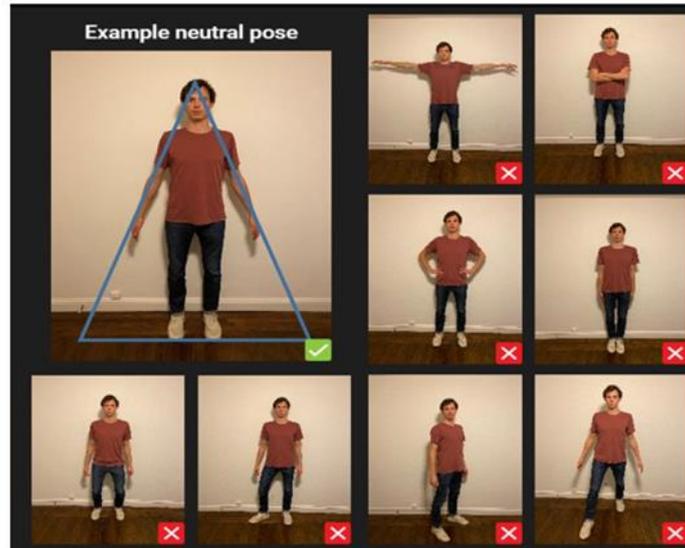


Figura 4.7. Pose neutral correcta y posiciones incorrectas.

Completada esta serie de acciones, ya es posible comenzar con la grabación de movimientos, los cuales, posteriormente, serán procesados automáticamente en la nube. Más adelante, será posible analizar esos datos, dado que OpenCap permite su descarga. La plataforma proporciona los archivos de videos, archivos *.mot* con los ángulos de las diferentes articulaciones y archivos *.trc* con la posición tridimensional de los marcadores anatómicos considerados.

5. CONCEPTOS SOBRE DEEP LEARNING

A continuación, se explicarán algunos de los conceptos básicos necesarios para comprender el *deep learning*, para acabar finalmente centrándonos en las redes convolucionales y recurrentes, las cuales son las arquitecturas más utilizadas a día de hoy en los estudios relacionados con el estudio de la marcha y su segmentación. Cabe mencionar que, para desarrollar el siguiente capítulo se ha utilizado como fuente de información los libros *Deep Learning* (Goodfellow et al., 2016) y *Deep Learning: principios y fundamentos* (Bosch Rué, 2019).

5.1 Introducción

Desde 1938, año en el cual Konrad Zuse creó la primera computadora, los ordenadores han experimentado un avance descomunal, llegando a sobrepasar límites, en términos de inteligencia, para ese entonces inimaginables. Así pues, actualmente definimos a la Inteligencia Artificial (IA) como la capacidad que poseen las máquinas de actuar como lo harían los seres humanos, siendo este un campo en pleno desarrollo y con aplicación en múltiples áreas, entre ellas la de la medicina y rehabilitación.

Se conoce como *machine learning* (aprendizaje automático) al subcampo de la IA, en el que las máquinas son entrenadas por personas para alcanzar su propio conocimiento a través del reconocimiento de patrones de datos. El desarrollo del machine learning supuso un gran logro ya que, gracias a esta innovación, los ordenadores empezaron a ser capaces de solucionar problemas del mundo real sin haber sido explícitamente programados para ello.

Durante las últimas décadas ha surgido una nueva vertiente dentro del *machine learning*, la cual es capaz llevar a cabo su función de forma prácticamente autónoma imitando la forma de trabajar del cerebro humano. Este tipo de tecnología más avanzada y sofisticada se denomina *deep learning* (aprendizaje profundo) y, normalmente, suele ser empleada para resolver problemas complejos que requieren de grandes cantidades de datos. Más concretamente, este subcampo del aprendizaje automático es capaz de generar, automáticamente, una jerarquía de conceptos, de forma que sobre la base de ideas simples va creando y definiendo elementos cada vez más complejos y abstractos. Hoy en día, es una técnica muy presente, utilizada por la mayor parte de empresas y compañías. Por ejemplo, Facebook hace uso del *deep learning* para analizar mensajes de conversaciones online o, centrándonos en el sector sanitario, es capaz de ayudar en el procesamiento de imágenes médicas y diagnóstico de enfermedades.

5.2 Conceptos básicos

Como se explicó anteriormente, el *deep learning* es un subcampo del aprendizaje automático. Por ello, para poder entender al primero es necesario aclarar y presentar una serie de conceptos básicos del segundo. Se dice que el machine learning se sustenta sobre tres pilares: los datos, las funciones matemáticas que constituyen el algoritmo y las métricas de rendimiento. Esto es así porque, tal y como apuntó Tom Mitchell en 1997 (Mitchell, 1997):

“Se dice que el algoritmo aprende de un determinado conjunto de datos, los cuales forman la experiencia E, para resolver una determinada tarea T, si el rendimiento del mismo al realizar la tarea T, medido a partir de métricas de eficiencia, mejora con la experiencia E”.

5.2.1 Tareas. Clasificación vs Regresión vs Agrupamiento

De entre las múltiples tareas que pueden ser ejecutadas por los algoritmos de *machine learning*, la clasificación y la regresión se podrían considerar como las principales.

En la clasificación, la tarea consiste en decidir a cuál de las posibles categorías de salida pertenece el dato dado a la entrada, estudiando para ello el conjunto de atributos que describen a este último. En este TFG la red abordará una tarea de clasificación. Concretamente, especificará a cuál de las posibles fases de la marcha pertenece el input de entrada. En cambio, las tareas de regresión predicen y proporcionan un valor numérico como dato de salida en vez de etiquetas de pertenencia a una clase.

Además, existen los algoritmos de agrupamiento, también denominados de *clustering*, que ante un conjunto de instancias proporcionadas deben crear, de forma automática, grupos en los que dividir a dichos datos. Así pues, en estos casos las categorías no están predefinidas, sino que el modelo tiene que generarlas basándose en patrones de similitud hallados en las entradas de la red.

Cabe mencionar que las métricas de rendimiento utilizadas variarán en función de la tarea. Por ejemplo, para problemas de clasificación suelen ser útiles las matrices de confusión o curvas ROC, mientras que en el caso de trabajos de regresión se prefieren otras métricas, como son el error absoluto medio o el error cuadrático medio.

5.2.2 Aprendizaje supervisado, no supervisado, semisupervisado y por refuerzo

Como se comentó anteriormente, los datos son una de las bases del aprendizaje automático. Esto es así porque constituyen la experiencia del algoritmo, necesaria para que este realice correctamente la tarea asignada. Según la forma en la que esos datos son dados podemos definir varios tipos de aprendizaje:

- **Aprendizaje supervisado.** El conjunto de datos de entrada está etiquetado, es decir, se conoce de antemano cuáles son sus valores de salida. De esta forma, se pretende que el algoritmo etiquete correctamente nuevos casos que no formaban parte del conjunto de entrenamiento. Para este trabajo se eligió este tipo de enfoque, haciendo uso de una interfaz para facilitar la tarea de etiquetado, la cual será presentada en el Capítulo 6. Materiales y Métodos.
- **Aprendizaje no supervisado.** Los datos no están etiquetados, es decir, desconocemos la clase a la que pertenecen cada uno de los mismos. Así pues, en este caso el objetivo consiste en que el sistema sea capaz de agrupar al conjunto de datos en grupos homogéneos pero lo más diferentes entre sí.
- **Aprendizaje semisupervisado.** Combina los dos tipos de aprendizajes anteriores, de modo que el conjunto de entrenamiento se compone de unos pocos datos etiquetados y una porción mayor sin etiquetar.
- **Aprendizaje por refuerzo.** En esta ocasión, lo que se busca es que el sistema pueda interactuar con el entorno, eligiendo la mejor estrategia para resolver una determinada tarea, siendo así su objetivo maximizar una “recompensa” o minimizar un “castigo”.

5.2.3 Overfitting y Underfitting

Tal y como su propio nombre expresa, el objetivo principal del *machine learning* es que el algoritmo sea capaz de aprender, a partir del entrenamiento, a realizar correctamente su tarea, sin haber sido explícitamente programado para ello. Esto es, que posteriormente consiga un buen rendimiento con un conjunto de datos totalmente nuevos. Esta capacidad de rendir adecuadamente ante futuras entradas desconocidas es lo que se denomina, durante la etapa de entrenamiento, generalización.

Así pues, durante el proceso de aprendizaje distinguimos dos conjuntos de datos principales: el set de entrenamiento y el set de test. El primero de ellos es usado durante el entrenamiento y el segundo para evaluar su capacidad de rendimiento y aprendizaje. Se dice que un algoritmo ha aprendido cuando tanto el error durante el entrenamiento como el cometido durante su posterior evaluación son pequeños y no existe una gran diferencia entre ambos.

Este concepto está estrechamente relacionado con el *overfitting* (sobreajuste) y el *underfitting* (subajuste). En casos de *overfitting*, el error de entrenamiento disminuye mientras que el de test aumenta poco a poco, es decir, el sistema no está aprendiendo la tarea, sino que únicamente está memorizando los datos de entrada formados por el set de entrenamiento. Por el contrario, durante el *underfitting* tanto el error de entrenamiento como el de test se encuentran elevados. Uno de los principales parámetros que afecta a estas ideas es la complejidad de nuestro modelo, de forma que cuando este es demasiado simple se observa la situación de subajuste, mientras que cuando es demasiado complicado se tiende al sobreajuste. Además, también depende de la calidad de los datos de entrenamiento, en otras palabras, que haya suficientes y, además, que las diferentes clases se encuentren balanceadas entre sí.

5.3 Redes neuronales

El *deep learning* es una forma de aprendizaje automático que se diferencia del resto por su manera de aprender, inspirándose en la organización y modo de funcionamiento del cerebro humano, utilizando para ello las denominadas redes neuronales artificiales.

Así pues, la neurona constituye la unidad básica de estos modelos. Estas se caracterizan según cuatro elementos principales: entrada (x), pesos (w), sesgo (b) y salida (y). A continuación, se muestra, de forma simplificada, el flujo de interacciones que tiene lugar en cada una de las neuronas de la red:

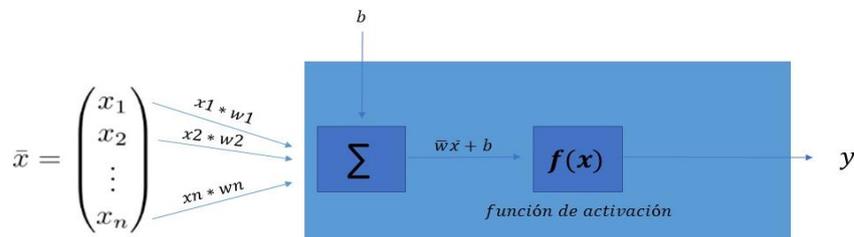


Figura 5.1. Esquema de una neurona artificial.

Tal y como se observa en la ilustración, la neurona toma un conjunto de datos de entrada, procedente de las conexiones con otros nodos, y los transforma en una salida. Para ello, debe aplicar las siguientes operaciones:

- **Transformación de entrada.** Multiplica al conjunto de entrada por sus respectivos pesos y le suma un sesgo. Matemáticamente, podría expresarse como $wx + b$.

- **Aplicación de la función de activación.** Modifica el valor transformado de entrada del nodo. Existen múltiples posibilidades, todas ellas no lineales. Entre las más comunes destacan la sigmoide, la tangente hiperbólica y ReLU. Esta última es capaz de resolver el problema de saturación que caracteriza a las otras dos, el cual será explicado después con mayor detalle.

Se considera que hablamos de *deep learning* cuando la red neuronal artificial está compuesta por más de tres capas, es decir, cuando tiene al menos una capa oculta. A mayor número de capas mayor será la profundidad de la red y, generalmente, mejor es su capacidad de predicción y poder de separación. En la siguiente figura se muestra un esquema de red neuronal profunda *feedforward*, primer y más simple tipo de red artificial creada.

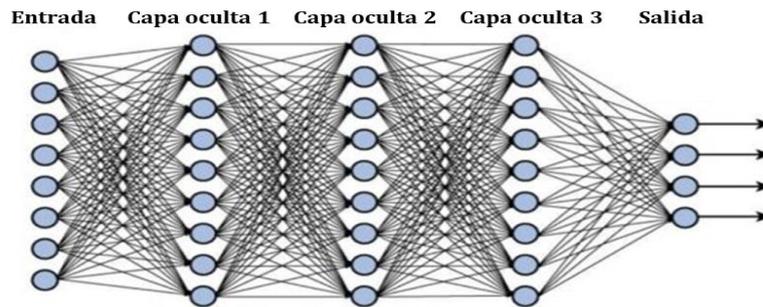


Figura 5.2. Arquitectura de una red neuronal profunda con múltiples capas (Parmar, 2018).

Brevemente, se podría resumir el proceso que tiene lugar durante el entrenamiento de la red de la siguiente manera. Cabe mencionar que los términos nombrados a continuación se explicarán con mayor profundidad más adelante. Así pues, los datos comienzan avanzando de izquierda a derecha, desde los nodos de entrada, pasando por las neuronas de las capas ocultas hasta las de salida (*forward propagation*). Una vez llegados al final, se compara el valor predicho con el valor real, obteniendo un resultado para la función de pérdidas. Dado que la red busca reducir el error cometido, aplicará el método del descenso de gradiente y el algoritmo de *backpropagation* para modificar los pesos de cada una de las capas y, de esta manera, conseguir adaptarse y aprender.

5.3.1 Función de pérdida

La función de pérdida, también denominada función de coste, es uno de los componentes básicos de nuestra red, dado que nos informa sobre lo lejos que se encuentra la predicción de salida de su correspondiente valor real. Existe un amplio abanico de funciones de pérdida, por ello es considerado un parámetro crítico a elegir durante el modelado de la red.

No obstante, la mayor parte de las veces se usa la entropía cruzada entre las etiquetas de entrenamiento y los valores predichos como función de coste. Su expresión matemática es la siguiente:

$$J(\theta) = - \sum_{i=1}^n t_i \log(p_i) \quad [1]$$

Para n clases, donde t_i es la clase verdadera y p_i la probabilidad de pertenencia a cada una de las clases predicha por la red.

Esta función es adecuada para modelos en los que el valor de salida es una probabilidad, como ocurre en los problemas de clasificación. Es por ello que esta ha sido la elegida como función de pérdidas en este trabajo de fin de grado.

5.3.2 Función de activación

Previamente se mencionó que la función de activación era necesaria para transformar la información de entrada en la salida, si bien, es mucho más importante que solo esto. Esta función es el mecanismo por el que las diferentes neuronas procesan y transmiten los datos a lo largo de la red, dado que, el valor obtenido a la salida de la misma tiene que ver con la fuerza de la señal, es decir, va a determinar si la neurona debe de activarse o no.

En los inicios del *deep learning*, la función sigmoide y tangente hiperbólica eran las más utilizadas. Ello se debe, fundamentalmente, a que ofrecen un valor acotado a su salida y a que su pendiente varía suavemente. Matemáticamente, pueden definirse como:

$$\text{Sigmoide}(z) = \frac{e^z}{e^z + 1} \quad [2] \quad \text{Tanh}(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \quad [3]$$

La función sigmoide toma cualquier valor de entrada y lo transforma en otro perteneciente al rango de 0 a 1, mientras que la salida de la tangente hiperbólica está acotada entre -1 y 1. A pesar de las ventajas comentadas unas líneas atrás, ambas presentan el problema de saturación. Esto es que, ante valores muy altos o muy bajos de entrada, la función se satura a sus respectivos extremos del rango. Este comportamiento se puede observar de forma mucho más clara cuando se las grafica.

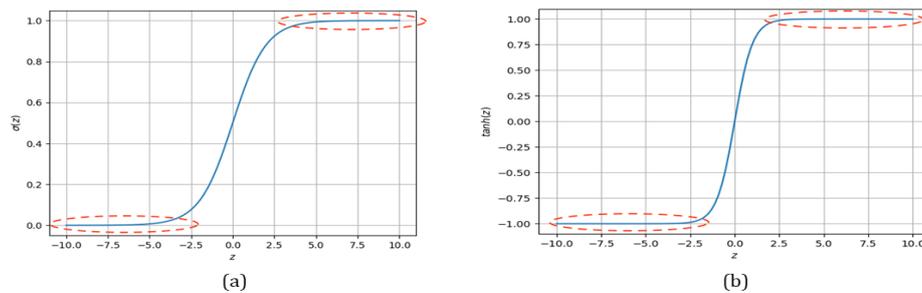


Figura 5.3. Funciones de activación. (a) Sigmoide. (b) Tangente hiperbólica. (*La Función de Activación*, s. f.).

En los 2000 apareció ReLU (*Rectifier Linear Unit*), otro tipo de función capaz de hacer frente al inconveniente de la saturación. Debido a ello, se ha convertido en la función de activación más utilizada en la actualidad y su expresión es la siguiente:

$$\text{ReLU}(z) = \begin{cases} 0, & z < 0 \\ z, & z \geq 0 \end{cases} \quad [4]$$

Es decir, genera una salida nula cuando la entrada es negativa, y un valor igual al de entrada cuando esta última es positiva. Su principal ventaja, además de su no saturación, reside en su velocidad de cálculo computacional, dado que su expresión matemática no es tan compleja como sucedía con las otras dos anteriores.

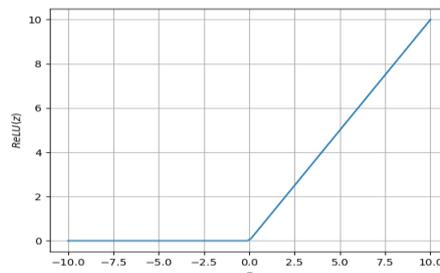


Figura 5.4. Función de activación ReLU (*La Función de Activación*, s. f.).

5.3.3 Gradiente Descendente

El descenso del gradiente es un algoritmo de optimización que constituye la base del entrenamiento en machine learning. En estos casos, va a ser utilizado para encontrar el mínimo de la función de pérdidas, definida y nombrada en apartados previos. Así pues, a partir del gradiente (o derivada) de esta función vamos a lograr el aprendizaje de nuestro modelo.

El gradiente descendente es un algoritmo iterativo, donde, en cada ciclo, se actualiza el valor de la variable considerada, en nuestro caso los pesos de la red, según se muestra en la siguiente ecuación:

$$w' = w - \mu \nabla J(\theta) \quad [5]$$

El parámetro μ hace referencia a la tasa de aprendizaje y su valor es crítico ya que influye en el aprendizaje de la red neuronal artificial. Valores muy pequeños de la misma hacen que se tarde mucho más tiempo en hallar el mínimo de la función, mientras que valores muy elevados provocan grandes cambios que dificultan el proceso de convergencia.

5.3.4 Backpropagation

Recapitulando lo explicado anteriormente, durante la fase de entrenamiento, la información de entrada fluye a lo largo de la red gracias al mecanismo de *forward propagation* hasta la salida, momento en el cual se genera una predicción. Al mismo tiempo se ha obtenido un valor para la función de pérdida $J(\theta)$, que hace referencia a la diferencia existente entre el dato predicho por el modelo y el real.

Una vez en este punto, el mecanismo de *backpropagation* hará ese mismo recorrido seguido, pero marcha atrás, actualizando el valor de los parámetros (según la operación matemática presentada en el apartado 5.2.3. Gradiente Descendente), intentando minimizar el error de la red y perfeccionando el proceso de predicción de la misma. Más concretamente, *backpropagation* trata de descubrir qué parte de carga tienen cada una de las neuronas en los errores cometidos, para posteriormente cambiar su configuración de pesos y sesgo. De esta manera, se logra corregir sus fallos y evitar, dada su influencia, los posibles errores en los nodos y capas siguientes.

5.4 Problema de desaparición del gradiente

Durante el entrenamiento de redes neuronales con múltiples capas ocultas suele darse un problema nombrado como la desaparición del gradiente. La responsabilidad de todo ello recae en las primeras capas, las cuales utilizan los términos de todas las neuronas posteriores para calcular su gradiente. Así pues, ante tanta profundidad de red se produce un estado intrínsecamente inestable, en el que las diferentes capas aprenden a distintas velocidades. Esta situación se denomina desaparición del gradiente, pudiéndose dar tanto una explosión del gradiente como una desaparición del mismo, dificultándose en ambas situaciones al aprendizaje de los primeros nodos de la red. Hoy en día, existen múltiples enfoques como las técnicas de regularización o las redes de memoria de largo y corto plazo (*Long Short Term Memory*, LSTM) que permiten lidiar con este inconveniente.

5.5 Optimización y Regularización

Las técnicas de optimización tratan de acelerar el aprendizaje de la red, así como mejorar su eficiencia, ayudando a minimizar los valores de la función de pérdida y aumentando su capacidad de generalización. Uno de los optimizadores más empleados en la actualidad, también utilizado durante este TFG, es el que se conoce como Adam. Este aplica una tasa de aprendizaje diferente a cada parámetro de la red, ajustándola según los valores obtenidos para el primer y segundo momento de gradientes. Estos momentos son calculados a partir del promedio de los gradientes pasados. De esta manera, logra una convergencia más rápida y evita los problemas de desaparición del gradiente.

Por otra parte, tal y como se comentó previamente, uno de los objetivos principales dentro del aprendizaje profundo es conseguir que la red aprenda a realizar la tarea sobre nuevos conjuntos de datos, es decir, que no se produzca *overfitting*. Se han desarrollado múltiples estrategias para evitar esta última situación, agrupadas bajo el nombre de técnicas de regularización. Dado que existe un gran número de técnicas, se explicarán las que han sido implementadas durante este trabajo:

- **Early Stopping.** Esta es la más popular de todas las posibles estrategias, debido a su simplicidad y efectividad. Básicamente, consiste en interrumpir el entrenamiento cuando el error en el set de validación comienza a incrementarse tras un determinado tiempo o número de épocas.
- **Reduce Learning Rate on Plateau.** Tal y como la anterior técnica, esta suele utilizarse muy frecuentemente. En este caso, al llegar a un punto conocido como *plateau*, donde las métricas medidas no aumentan o incluso empeoran, se disminuye el valor de la tasa de aprendizaje. Para su implementación debe fijarse un valor de paciencia, que se refiere al número de épocas en las que se permite que la métrica seleccionada no mejore.
- **Dropout.** Esta técnica va a desactivar de forma temporal y aleatoriamente algunas de las neuronas que conforman las capas ocultas de nuestra red. Como consecuencia, estas no formarán parte de los cálculos que tienen lugar durante forward propagation y backwardpropagation, disminuyendo así la dependencia que ejercen sobre las neuronas vecinas. *Dropout* se caracteriza por un parámetro, el cual hace referencia a la probabilidad que tiene una neurona de ser desactivada, pudiendo tomar un valor entre 0 y 1.
- **Batch Normalization.** Esta estrategia consiste en normalizar las activaciones de cada capa, de forma que tengan una media próxima a cero y una desviación estándar cercana a uno. Gracias a esta operación, se logra disminuir la covarianza interna y se facilita la convergencia de la red.

5.6 Redes neuronales convolucionales

Entre los distintos tipos de modelos de redes neuronales artificiales encontramos las redes convolucionales, las cuales se asemejan a la vía de la visión humana. Su éxito y fuerza recaen en la operación de convolución, tal y como su propio nombre indica, dado que este método les permite disminuir la dimensionalidad de los datos al mismo tiempo que conservan el máximo de información de estos.

Otra de las virtudes de las redes convolucionales reside en la compartición de parámetros. Más concretamente, esto quiere decir que todas las neuronas ocultas pertenecientes a la misma capa utilizan los mismos valores de pesos y sesgo. El motivo de ello recae en que, en cada capa, los diferentes nodos identifican y evalúan la misma característica, solo que cada uno lo hace en lugares diferentes del dato de entrada. De este modo, los diferentes filtros se definen y caracterizan según ese conjunto de pesos y sesgos compartidos. Así pues, el reparto de pesos y sesgo supone disminuir en grandes cantidades el número de parámetros de la CNN. Asimismo, esto deriva en un entrenamiento mucho más rápido y en una red mucho menos compleja y más fácil de comprender. Por último, gracias a la compartición y al menor número de parámetros necesarios será posible construir redes de mayor profundidad sin que se dispare el coste computacional de la misma.

Dadas estas capacidades, hoy por hoy, este tipo de redes se emplean en una gran variedad de ámbitos, desde el procesamiento de imágenes y vídeos hasta el desarrollo de vehículos autónomos, dado que son capaces de trabajar tanto con entradas de dos o tres dimensiones como con series temporales.

5.6.1 La operación de convolución

La convolución es una operación que transforma dos funciones ($f(t)$ y $h(t)$) en una tercera ($s(t)$), la cual muestra como una de ellas es modificada por la otra. Para señalarla se suele emplear un asterisco (*), tal y como se enseña a continuación:

$$s(t) = (f * h)(t) = \int_{-\infty}^{\infty} f(\tau)h(t - \tau)d\tau \quad [6]$$

Este operador puede ser empleado en distintos campos, como la probabilidad, estadística, física, óptica... Ahora bien, centrándonos en las redes convolucionales, la función $f(t)$ representa al dato de entrada, mientras que $h(t)$ se refiere al filtro, también conocido como *kernel*. Por otra parte, la salida $s(t)$ se suele denominar mapa de características o *feature map* en inglés. En la Figura 5.5 se ilustra un ejemplo de convolución sobre un dato de entrada bidimensional (*input*), como podría ser una imagen. Tal y como puede verse, el *kernel* únicamente interactúa en las posiciones en las que cuadra con el *input*, es decir, no llega a salirse del borde. Sin embargo, a continuación, se explicará que no siempre esto tiene por qué ser así.

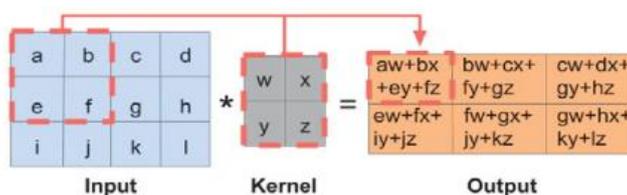


Figura 5.5. Ejemplo de convolución sobre una entrada en 2D (Goodfellow et al., 2016).

5.6.2 Filtro o Kernel

En las CNN denominamos *kernels* a las matrices de pesos que encontramos en cada una de las capas convolucionales. Estos pesos constituyen el conjunto de parámetros que la red debe de ajustar durante su entrenamiento.

La entrada, al igual que el filtro, es un elemento caracterizado según tres dimensiones: alto, ancho y número de canales. Así pues, el *kernel* debe coincidir en la última de estas

dimensiones con el input, mientras que las otras dos restantes son a gusto del diseñador, si bien suele tener un tamaño mucho más pequeño que el dato de entrada. Por ejemplo, con imágenes a color se tienen tres canales de entrada, por lo que el filtro tendrá el mismo número en dicha dimensión, siendo la salida correspondiente $(R * Kernel_{1,1}) + (G * Kernel_{1,2}) + (B * Kernel_{1,3})$. Luego, tal y como puede deducirse de la expresión anterior, se generarán tantos mapas de salida como *kernels* sean usados, teniendo estos un tamaño más pequeño que su respectiva entrada (excepto con filtros de dimensión 1x1), dada la operación de convolución. De esta forma, se logra esa reducción de la dimensionalidad, una de las ventajas principales de este tipo de redes.

5.6.3 Padding

Es cierto que, si queremos obtener una red neural de gran profundidad, tal vez no nos interese reducir tan rápido la dimensionalidad de los datos en las primeras capas de la misma. No obstante, para hacer frente a este problema existe un concepto denominado *padding*. Básicamente consiste en añadir tantos ceros como sean necesarios alrededor de la entrada, hasta conseguir que esta y su respectiva salida tengan el mismo tamaño. Este efecto puede verse de forma más clara en la Figura 5.6, donde se muestra la dimensión del mapa de características con y sin *padding*.

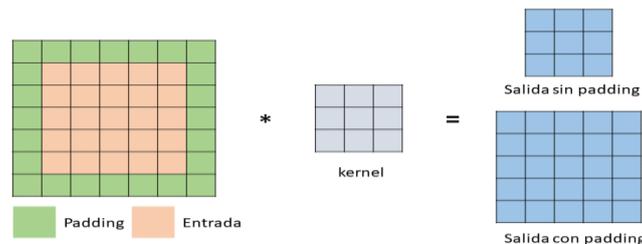


Figura 5.6. Efectos del padding en la dimensionalidad de salida

5.6.4 Capa de agrupamiento (*pooling*)

Normalmente, la estructura típica de una CNN consta de tres niveles. El primero de ellos estaría formado por la operación de convolución, cuyo resultado final consiste en el conjunto de características extraídas. En la segunda capa los *feature map* pasan por una función de activación, generalmente la función ReLU, con intención de introducir no linealidad a los datos. Por último, es habitual introducir en el tercer nivel la función de *pooling*, cuyo fin es hacer más robustas a algunas de las características. A su vez, consiguen reducir el tamaño de los mapas de salida y prevenir, de este modo, el *overfitting*.

Existen varias formas de realizar esta operación, siendo el *max-pooling* la más utilizada. Sencillamente, esta última consiste en seleccionar el valor máximo de cada una de las regiones formadas en el mapa de características. Además, también es conocida otro tipo de agrupamiento, el *average-pooling*, la cual hace el promedio de todos los valores del conjunto.

5.7 Redes neuronales recurrentes

Las redes neuronales recurrentes constituyen un tipo de arquitectura muy empleada en *deep learning* para la resolución de problemas secuenciales. Luego, resultan de gran utilidad para analizar los patrones de marcha, dado que estos constan de una secuencia de movimientos correlacionados entre sí.

A pesar de todas las ventajas explicadas con anterioridad sobre las CNN, uno de los inconvenientes de las mismas radica en su rigidez, en cuanto a su necesidad de que los datos de entrada y salida siempre tengan las mismas dimensiones. No obstante, las secuencias y videos analizados no siempre van a tener el mismo número de muestras y *frames*, respectivamente. Además, las redes convolucionales no son adecuadas para determinar la relación entre los datos de una secuencia. Sin embargo, las redes recurrentes no presentan estas dos limitaciones, lo que las convierte en un tipo de estructura también muy popular.

Estas capacidades radican en su ‘memoria’, dado que al generar una salida (también denominada activación) ante un dato de entrada, la red no solo tiene en cuenta a esta última, sino que también usa la activación obtenida en la iteración anterior. Así pues, dado este gran potencial no es de extrañar que actualmente estén presentes en tan diversas aplicaciones, desde el traductor de Google hasta el asistente de voz personal Siri.

5.7.1 Estructura de las redes recurrentes

Como se mencionó atrás, el poder de este tipo de redes viene dado por su memoria. A continuación, al mostrar gráficamente su estructura, este concepto quedará mucho más claro, pero, básicamente, esta facultad se debe a que existen dos entradas y dos salidas para cada instante temporal.

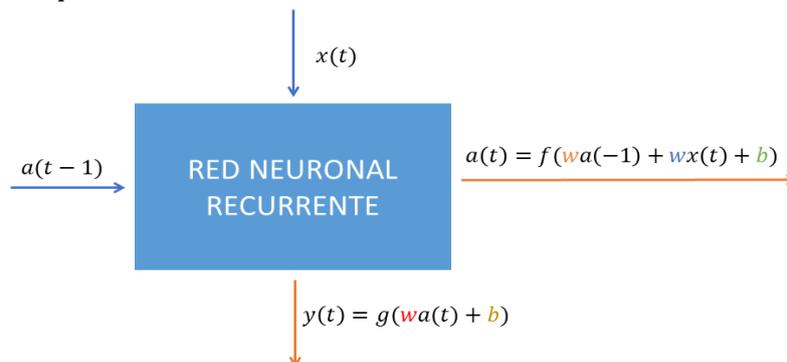


Figura 5.7. Estructura básica de una red neuronal recurrente.

Prestando atención a la ilustración, vemos que las entradas son el nuevo dato ($x(t)$) y la activación anterior ($a(t-1)$), en tanto que las salidas son la predicción ($y(t)$) y la activación ($a(t)$), también conocida como *hidden state*, para ese instante de tiempo. También, de la figura podemos deducir el modo en el que se calculan ambas salidas, siendo este igual al ya visto para el resto de redes neuronales, es decir, transformando los datos y aplicando posteriormente una función de activación. Sin embargo, mientras que para calcular $a(t)$ se hace uso de las dos entradas, para conseguir $y(t)$ se precisa esta obtenida anteriormente, dicho en otras palabras, el *hidden state* actual. Asimismo, los pesos y *bias* constituyen los parámetros del modelo, los cuales serán determinados a lo largo de la etapa de entrenamiento.

De este modo, vemos que la retroalimentación que se da al depender $y(t)$ de $a(t)$, que a su vez depende de $a(t-1)$, es lo que permite dotar de temporalidad a la red, consiguiendo retener y compartir los datos entre instantes sucesivos.

6. MATERIALES Y MÉTODOS

Durante este capítulo se explicarán de forma detallada todas las acciones llevadas a cabo y materiales necesarios para la elaboración de este trabajo de fin de grado, desde la adquisición de la base de datos, pasando por el procesado de los mismos, hasta las características del entrenamiento de las redes *deep learning* creadas bajo el objetivo de segmentar la marcha en varias de sus fases principales.

6.1 Base de datos

Para la ejecución de este proyecto se adquirió una base de datos del movimiento de marcha. Esta adquisición de datos se realizó empleando simultáneamente tecnología de captura de video y sensores vestibles. Los sujetos de estudio fueron voluntarios sanos, tanto adultos como niños, y pacientes pediátricos diagnosticados de parálisis cerebral. Los adultos fueron grabados en las instalaciones de la ETSI de Telecomunicaciones de la Universidad de Valladolid y los niños en el Hospital Universitario Río Hortega de Valladolid.

6.1.1 Protocolo de adquisición de la base de datos

La adquisición de la base de datos se realizó en la sala “Leonardo Torres Quevedo” de la ETSI de Telecomunicación de la Universidad de Valladolid y en la sala de rehabilitación del Hospital Universitario Río Hortega. A continuación, se muestra la disposición de las salas de captura:

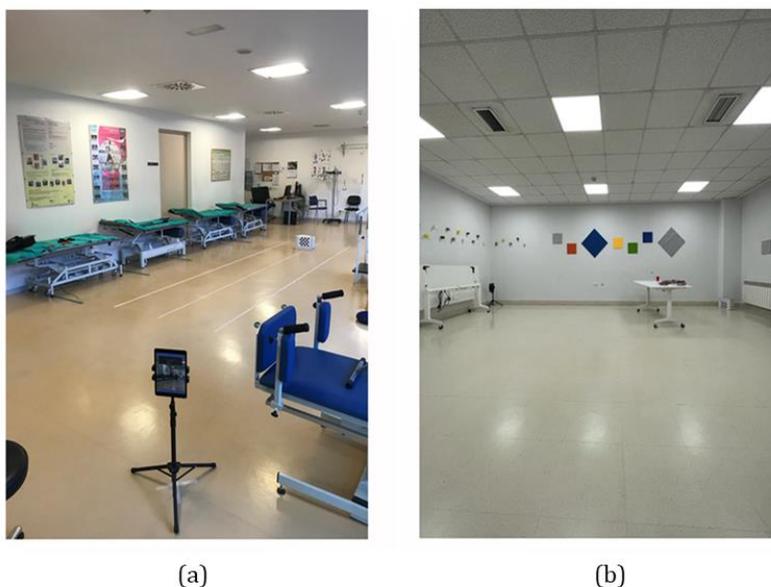


Figura 6.1. Disposición de las salas empleadas para la captura de datos. (a) Sala de rehabilitación del Hospital Universitario Río Hortega. (b) Sala “Leonardo Torres Quevedo”.

Para la adquisición de muestras con sensores vestibles se hizo uso de los dispositivos de la Figura 6.2. Concretamente, se corresponden con las unidades de medición inercial Xsens MVN Awinda, las cuales se fijan a las extremidades del sujeto de interés a partir de cintas de velcro, tal y como se muestra en la Figura 6.2. Se siguieron las recomendaciones especificadas en la Tabla 6.1. No obstante, cabe destacar que al tratarse de IMU, la precisión en la posición de los sensores no es tan importante, ya que estos miden orientaciones y posiciones. De este modo, se usaron un total de 17 IMU para el estudio de movimientos de:

foot (2), lower-legs (2), upper-legs (2), pelvis (1), hands (2), lower-arms (2), upper-arms (2), shoulders (2), chest (1), head (1). No existe contacto entre los dispositivos y la piel.

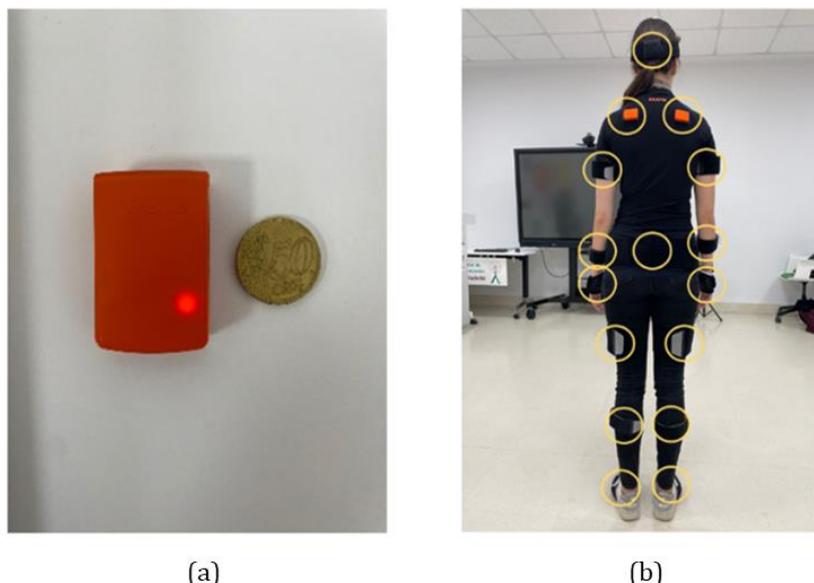


Figura 6.2. Adquisición de datos con sensores vestibles. (a) Sensor Xsens MVN Awinda. (b) Disposición de los sensores en el cuerpo del sujeto mediante las cintas de velcro.

Sensores	Recomendación
Shoulders	Posicionar estos sensores en horizontal, para lograr una mejor fijación de los mismos a la camiseta.
Lower-arms	Colocar los sensores lo más próximo de la muñeca posible, para conseguir más información sobre el movimiento de pronación y supinación.
Pelvis	Este sensor debe estar en posición vertical, de forma que la palabra 'Xsens' se encuentre en la parte superior.
Upper-legs	Colocar estos sensores en el centro del muslo lateral, para evitar el tejido graso y que el velcro caiga hacia abajo.

Tabla 6.1. Recomendaciones sobre la colocación de los dispositivos Xsens MVN Awinda.

En cuanto a la captura de datos mediante los sistemas de vídeo, se hizo uso de 2 Apple iPad® y la plataforma de código abierto OpenCap. Estos primeros estaban situados en puntos estratégicos para lograr una visión completa del cuerpo del sujeto durante el desarrollo de las actividades. Además, para conseguir que ambos se encontrasen siempre en la misma localización durante todas las sesiones, se colocaron unas marcas en el suelo.

- La captura de la base de datos de sujetos sanos se realizó en la ETSI de Telecomunicación de la Universidad de Valladolid, de acuerdo al protocolo con referencia PI 21-2341 v3.0, Adquisición de base de datos de movimientos del cuerpo humano en actividades de la vida diaria empleando vídeo y sensores vestibles, que recibió el dictamen favorable del comité de ética de la investigación con medicamentos CEIm ÁREA DE SALUD VALLADOLID ESTE en febrero de 2023.
- La captura de la base de datos de paciente pediátrico se realizó en el Hospital Universitario Río Hortega, de acuerdo al protocolo con referencia 23-PI038, que recibió el dictamen favorable del comité de ética de la investigación con medicamentos CEIm ÁREA DE SALUD VALLADOLID OESTE en marzo de 2023.

Con cada uno de los voluntarios, se procedió de acuerdo con el siguiente protocolo:

1. Con anterioridad a la llegada del sujeto se comprobó el estado de batería de los sensores, para asegurarse de que estos tuvieran suficiente carga. Además, se verificó la correcta disposición de la sala, prestando especial interés en que los dispositivos de captura de video estuvieran colocados sobre sus correspondientes marcas y que no hubiese ningún tipo de oclusión.
2. Se realizó una breve explicación del estudio y se facilitó el consentimiento informado, para que el sujeto pudiese leerlo con atención y lo firmase.
3. Una vez firmado el consentimiento, se le asignó un número de identificación, para respetar su privacidad, y se recogió información del mismo a mayores: edad (años), sexo, altura (m), peso (kg) y número de pie. En el caso de los sujetos pediátricos con PC también se recogió el tipo de afectación que padecían.
4. A continuación, se procedió a colocar los sensores vestibles de acuerdo con la Figura 6.2. Para la calibración de los dispositivos IMU se siguieron las instrucciones marcadas por Xsens MVN Awinda. Después, se continuó con el calibrado de los sistemas de vídeo. Ambos pasos fueron explicados con anterioridad en el Capítulo 4. Análisis Instrumentado de la Marcha.
5. Posteriormente, se pidió al sujeto el desarrollo de la actividad de marcha descrita en la Tabla 6.1. Se realizaron un total de 2 tomas, para adelantarse ante posibles errores en alguna de las grabaciones y evitar tener que llamar al sujeto de nuevo. En el caso de que las dos tomas resultasen válidas, ambas fueron utilizadas como datos para el desarrollo de la red neuronal. En ambas tomas los sujetos pediátricos caminaron sin zapatos, con el objetivo de poder observar más claramente el movimiento del pie. Además, en aquellos sujetos pediátricos con parálisis cerebral que portaban una férula multiposicional pie-tobillo, se tomó una tercera toma con este dispositivo de ayuda para la marcha.
6. Una vez realizadas las capturas, se retiraron los dispositivos vestibles. Todos los archivos generados durante la grabación se guardaron en unas carpetas que tienen como nombre el identificador del sujeto.

Tal y como se comentó anteriormente, los dispositivos Xsens MVN Awinda permitían hacer su propio análisis de la marcha en MotionCloud, siempre y cuando se diesen como mínimo 10 pasos. Desafortunadamente, el reducido espacio del que se disponía en la sala Leonardo Torres Quevedo, hizo inviable cumplir con ese requisito. No obstante, en el Hospital Universitario Río Hortega se pudo realizar una captura adicional a los sujetos pediátricos, la cual se denominó 'Marcha Larga', descrita en la Tabla 6.1. Es necesario mencionar que, en esta última grabación nombrada, únicamente se emplearon los sensores vestibles y no OpenCap. La causa de ello fue que una marcha con tantos pasos se salía de la zona de visión de las cámaras, haciendo inviable la adquisición de datos con sistemas basados en vídeo como lo es OpenCap.

Actividad	Número de tomas	Instrucciones para el sujeto
Marcha	2	"Sitúese sobre la marca del suelo, mirando hacia otra marca colocada enfrente de usted. A continuación, camine ida y vuelta en línea recta siguiendo el camino delimitado por ambas marcas, dando siempre el primer paso del camino de ida y de vuelta con el pie derecho."

Marcha con férula	1	"Sitúese sobre la marca del suelo, mirando hacia otra marca colocada enfrente de usted. A continuación, camine ida y vuelta en línea recta siguiendo el camino delimitado por ambas marcas, dando siempre el primer paso del camino de ida y de vuelta con el pie derecho."
Marcha larga	1	"Sitúese sobre la marca del suelo, mirando hacia otra marca colocada enfrente de usted. A continuación, camine 12 pasos ida y 12 pasos vuelta en línea recta, dando siempre el primer paso del camino de ida y de vuelta con el pie derecho."

Tabla 6.2. Descripción de la actividad a realizar durante la captura de datos.

6.1.2 Descripción de los sujetos de base de datos

El *dataset* adquirido y utilizado consta de un total de 25 sujetos, de los cuales 17 son de control, es decir, no tenían ningún tipo de deterioro en la marcha, y los 8 restantes niños con parálisis cerebral. Más concretamente, dentro de los primeros, 14 son adultos y 3 sujetos pediátricos. Sus características se muestran en las siguientes tablas:

Identificador	Edad	Sexo	Altura	Peso	Nº Pie
S01	21	H	1.89	72	42.5
S02	21	M	1.59	66	38
S03	21	H	1.72	68	41.5
S04	21	H	1.82	84	43
S05	21	H	1.78	70	42
S06	21	H	1.80	74	43
S07	21	M	1.65	68	38
S08	21	M	1.74	55	39
S09	21	M	1.68	60	39
S10	21	H	1.73	80	43
S11	11	H	1.52	35	39
S12	20	M	1.60	48	38
S13	21	M	1.71	67	40
S14	22	M	1.53	50	37
S15	21	H	1.80	60	42
S16	7	H	1.27	23	29
S17	9	H	1.38	29	34

Tabla 6.3. Características de los sujetos de control de la base de datos.

Identificador	Edad	Sexo	Altura	Peso	Nº Pie	Afectación
P01	11	H	1.51	38	39.5	Hemiparesia Derecha
P02	6	H	1.16	18	29	Hemiparesia Izquierda
P03	7	H	1.16	18	29	Hemiparesia Izquierda
P04	8	H	1.29	28	33	Hemiparesia Derecha
P05	8	H	1.30	31	39	Paraplejia
P06	13	H	1.64	57	41	Hemiparesia Izquierda
P07	5	H	1.11	20	29	Hemiparesia Izquierda
P08	9	H	1.35	27	34	Hemiparesia Izquierda

Tabla 6.4. Características de los sujetos pediátricos con PC de la base de datos.

6.2 Etiquetado de los datos

Para el entrenamiento de la red neuronal se siguió un enfoque de aprendizaje supervisado. Así pues, para facilitar el etiquetado de datos se desarrolló la siguiente interfaz en Python:



Figura 6.3. Interfaz desarrollada en Python para el etiquetado de datos.

Tal y como puede verse en la Figura 6.3, en la interfaz es posible visualizar cuatro imágenes. Estas se corresponden a los *frames* registrados por el sistema OpenCap. Concretamente, las dos primeras imágenes proceden del vídeo capturado por una de las cámaras, mientras que las otras dos restantes pertenecen al otro dispositivo de grabación. Además, hay una serie de *sliders* que permiten que se muestren en esas pantallas los *frames* deseados, ya que, tienen una longitud igual al número de muestras de los videos importados. Por último, es necesario comentar que, al abrir ambos videos y, sabiendo que estos tienen la misma duración y *fps*, se generan automáticamente dos vectores de ceros de igual longitud al número de muestras de dichos vídeos, una para pie izquierdo y otra para pie derecho. Estos vectores, durante la tarea de etiquetado, irán siendo modificados al asociarles un número concreto en función de la fase de la marcha a la que se corresponda la respectiva muestra del vídeo a la que representan.

Así pues, el procedimiento de etiquetado fue el siguiente:

- Con el *slider* superior se seleccionaba el *frame* que se correspondía con el inicio de una fase de la marcha. Esta imagen era mostrada en la primera y tercera pantalla y así, de este modo, se conseguía una visión desde dos puntos evitando posibles oclusiones de los pies que afectasen a la precisión y objetividad de esta tarea.
- Con el *slider* inferior se indicaba qué muestra se asociaba con el final de dicha fase. Al igual que ocurría antes, este *frame* era mostrado en la segunda y cuarta pantalla para tener dos vistas distintas.
- Una vez teniendo claro dichos *frames* de inicio y final, se pulsaba uno de los botones ubicados en el cuadro inferior derecho. De esta manera, todas las muestras ubicadas entre ese rango de inicio y final quedaban descritas por el número asociado a esa fase.

Una vez terminada la tarea de etiquetado, cada muestra quedó caracterizada por una de las siguientes etiquetas:

- **0 (No movimiento).** No se corresponde con ninguna fase de la marcha. Se decidió utilizar una clase de etiqueta para esta situación dado que, al inicializar las grabaciones, puede darse un retardo desde que se le pide al sujeto que comience a caminar hasta que este hace dicha acción. Asimismo, también puede haber un retraso desde que el sujeto termina de realizar la actividad hasta que se pausa la grabación. Por ello, existen momentos en los cuales no hay movimiento a los que se les asoció este número. Además, dado el limitado espacio de la sala, al llegar a cada uno de los extremos marcados había ocasiones en las que el sujeto no llegaba a completar del todo el movimiento propio de alguna de las fases de la marcha, por lo que a esas muestras incompletas también se las caracterizó con la etiqueta 0.
- **1 (Apoyo inicial).** Corresponde con lo que se ha denominado como fase inicial de apoyo. Esta integra tanto el contacto inicial como la respuesta a la carga, es decir, va desde que el talón del pie de referencia toca el suelo hasta que el miembro contralateral está a punto de elevarse y dejar de estar en contacto con el suelo.
- **2 (Apoyo medio).** Esta etiqueta se asocia con la fase media de apoyo, dicho de otra forma, cuando todo el peso recae sobre el pie de referencia al haberse levantado del suelo el otro.
- **3 (Apoyo final).** Este número une a la fase final de apoyo y *foot off*, por lo tanto, caracteriza a aquellas muestras que van desde que el pie contralateral hace contacto con el suelo mediante el talón hasta que los dedos del pie de referencia abandonan el piso. En este trabajo, estas dos etapas se han reunido bajo el nombre de fase final de apoyo.
- **4 (Swing).** Por último, la fase completa de balanceo o swing se asociará con esta etiqueta.

Los límites de cada una de estas fases de movimiento consideradas son expuestos de forma más visual a continuación, con la finalidad de lograr un mayor entendimiento de las mismas:

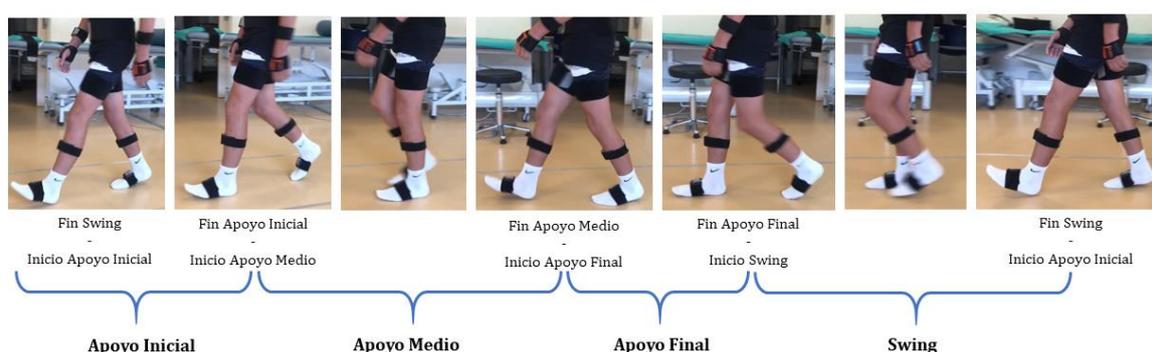


Figura 6.4. Fases de la marcha (1-4) correspondientes al pie izquierdo.

Asimismo, al inspeccionar la base de datos de sujetos pediátricos, se observó que, en algunas de las grabaciones, típicamente aquellas correspondientes a las tomas sin la férula multiposicional, los sujetos no realizaban correctamente ciertos movimientos. Concretamente, el fallo se daba al finalizar la fase de swing y comenzar la de apoyo inicial, dado que estos pacientes no realizaban la flexión dorsal típica de este momento de la marcha, pisando el suelo bien con la punta o bien con toda la superficie del pie a la vez. Por ello, se decidió añadir dos etiquetas a mayores, las cuales fueron usadas en algunos de los

entrenamientos. Comentar que, con la intención de asegurar que realmente se estaba dando una pisada en punta, se comprobó el parámetro ‘Contador de eventos de contacto’ que el análisis de MotionCloud asignaba al fichero ‘.mvt’ correspondiente a esa misma captura de marcha:

- **5 (Swing Parálisis Cerebral).** Este número se corresponde con una fase de swing anómala, más específicamente, cuando el pie de referencia se dispone a finalizar esta fase sin encontrarse realizando una flexión dorsal.
- **6 (Apoyo inicial Parálisis Cerebral).** Etiqueta asociada con un inicio patológico del apoyo inicial, donde, al igual que ocurría con la fase explicada anteriormente, el pie se encuentra completamente apoyado en el suelo o sobre su punta, en vez de comenzar en dorsiflexión.

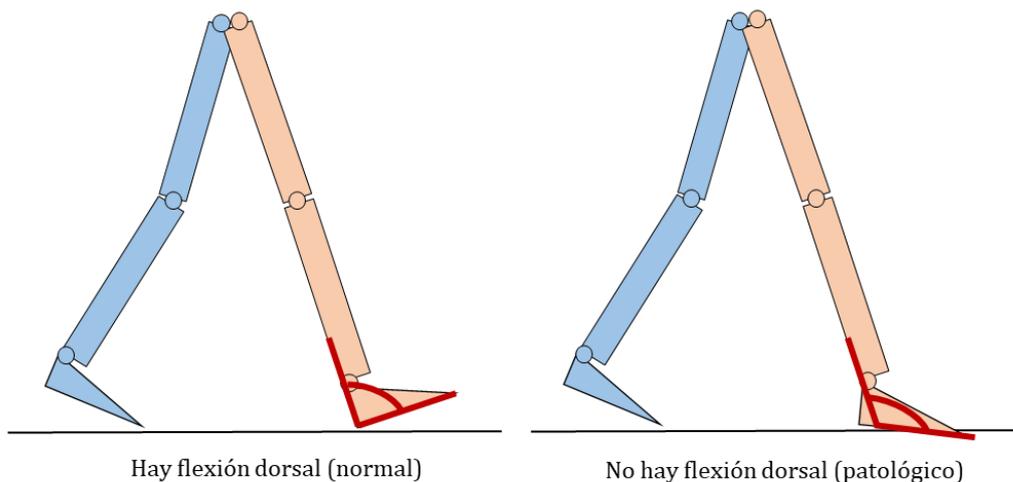


Figura 6.5. Representación de la dorsiflexión que se dio en sujetos normales y patológicos.

Así pues, se han empleado siete etiquetas donde realmente, se representan únicamente cuatro momentos de la marcha. Es cierto que, en trabajos como el de Kreuzer & Munz (2021), diferenciaban un total de cinco fases distintas, así como, tal y como se explicó en el Capítulo 3. Proceso de la Marcha, se distinguen un total de ocho posibles etapas de la marcha. Sin embargo, para este TFG se ha decidido emplear este número de etiquetas ya que, al trabajar con sujetos no solo sanos, sino también patológicos, esas distintas ocho fases al caminar pueden no llegarse a dar correctamente en estos últimos, lo que haría de la tarea manual de etiquetado algo muy subjetivo y no exacto del todo que podría perjudicar al rendimiento y veracidad del modelo neuronal.

6.3 Pre-procesado de la base de datos

Antes de poder usar la base de datos adquirida como conjunto de entrenamiento, esta tuvo que someterse a una serie de modificaciones.

Como paso previo a la etapa de pre-procesado como tal, generamos conjuntos de muestras en función de su etiqueta y pie de referencia. Explicado con más detalle, esto quiere decir que, para cada una de las dos tomas de los distintos sujetos, se agruparon en un mismo conjunto aquellas muestras que compartían el mismo número de etiqueta, y todo ello, tanto para pie izquierdo como derecho. De esta manera, ahora cada uno de los sujetos queda caracterizado un total de 20 archivos, mientras que al comienzo tan sólo se disponía de 2

(las dos tomas capturadas). Por ejemplo, para el primer sujeto, los 10 archivos generados para el miembro izquierdo fueron:

Trial/ Etiqueta	0	1	2	3	4
1	S01-L00-T01	S01-L01-T01	S01-L02-T01	S01-L03-T01	S01-L04-T01
2	S01-L00-T02	S01-L01-T02	S01-L02-T02	S01-L03-T02	S01-L04-T02

Tabla 6.5. Archivos del lado izquierdo generados como paso previo al pre-procesado.

Todo ello se repetiría para el lado derecho. Así pues, a partir de esta tabla explicativa puede intuirse la notación elegida para nombrar los archivos:

- **Sx.** Donde la x indica el número de sujeto. Para nuestra base de datos va de 01-15. En el caso de que el individuo se corresponda con uno patológico Sx se cambia por Px.
- **L0x.** en este caso, la x hace referencia al número de etiqueta, pudiendo tomar valores del 0-4. Concretamente, la letra L es la asignada al lado izquierdo, sin embargo, si se tratase del pie derecho el archivo se identificaría con una R. Asimismo, es necesario comentar que, en aquellos sujetos patológicos en los que se consideró necesario etiquetar ciertos *frames* con los números 5 y 6, existen archivos con valor Px-L05-Tx y Px-L06-Tx (lo mismo para el miembro derecho si se diese el caso).
- **Tx.** esta última notación muestra el número de toma, cuyo valor de x puede ser 01, 02 o 03, este último en el caso de pacientes con PC que tuvieron que hacer una toma a mayores sin férula multiposicional.

Este primer paso se continúa con el pre-procesado. Es importante apuntar que, en esta etapa, no se utilizaron aquellos archivos correspondientes a la etiqueta 0 (L00 o R00), es decir, aquella que se asociaba con la ausencia de movimientos o cuando estos eran incompletos. Se decidió llevar a cabo esta acción dado que, consideramos que esos datos podrían confundir a la red neuronal en su tarea de clasificación, dificultando su aprendizaje y limitando su rendimiento.

Esta etapa se caracteriza a su vez por varias operaciones. Dado que, en cada uno de los diferentes entrenamientos ejecutados, bien no todas estas se realizaron o bien se usaron valores diferentes en cada uno de ellos, a continuación, se explicará de forma generalizada qué hacen cada una de esas operaciones:

- **Selección de características.** Es posible elegir qué columnas nos interesan de la totalidad de datos proporcionados por los archivos. Estas características se corresponden con ángulos de las articulaciones del cuerpo. Por ejemplo, el conjunto de los 35 ángulos recogido en los ficheros '*mot*' de OpenCap viene resumido en la Tabla 6.5. Uno de los objetivos del entrenamiento fue comprobar si para obtener un buen rendimiento bastaba únicamente con los ángulos correspondientes a la parte inferior del cuerpo o, si por el contrario este mejoraba al añadir datos de los miembros superiores. Además, se trató de comprobar qué características eran las que contenían la mayor parte de la información, de modo que en el modelo neuronal

se utilizase el menor número de ellas para reducir el coste computacional y evitar el *overfitting*.

- **Segmentación.** Cada uno de los archivos obtenidos se segmentan en conjuntos más pequeños de datos, de tal forma que aumentamos el *dataset* disponible. Para ello se hizo uso de una ventana deslizante, cuya longitud y porcentaje de solapamiento fue variando en los distintos entrenamientos.

Ángulos miembro inferior	pelvis_tilt, pelvis_list, pelvis_rotation, pelvis_tx, pelvis_ty, pelvis_tz, hip_flexion_r, hip_adduction_r, hip_rotation_r, knee_angle_r, knee_angle_r_beta, ankle_angle_r, subtalar_angle_r, mtp_angle_r, hip_flexion_l, hip_adduction_l, hip_rotation_l, knee_angle_l, knee_angle_l_beta
Ángulos miembro superior	ankle_angle_l, subtalar_angle_l, mtp_angle_l, lumbar_extension, lumbar_bending, lumbar_rotation, arm_flex_r, arm_add_r, arm_rot_r, elbow_flex_r, pro_sup_r, arm_flex_l, arm_add_l, arm_rot_l, elbow_flex_l, pro_sup_l

Tabla 6.6. Variables angulares proporcionadas por OpenCap.

6.4 Red neuronal profunda para la segmentación de la marcha

Durante este apartado, se expondrán las distintas arquitecturas de redes neuronales desarrolladas para llevar a cabo la tarea de clasificación. Para su elaboración se hizo uso de Keras, una biblioteca de código abierto de alto nivel dedicada a la creación de redes neuronales (*Keras: Deep Learning for humans*, s. f.).

- DNN

La primera de las redes neuronales desarrollada para este estudio fue una *fully-connected*, que, tal y como se explicó en el Capítulo 5. Conceptos sobre Deep Learning, consiste en una estructura donde todas las neuronas están totalmente conectadas entre capa y capa. El número de neuronas de la capa de entrada viene determinado por las dimensiones del dato de entrada, es decir, por el número de filas (número de muestras), columnas (número de variables) y canales. Para este trabajo en concreto, el número de canales siempre tuvo un valor igual a la unidad, sin embargo, el número de muestras y variables fue modificándose a lo largo de los distintos entrenamientos, a fin de optimizar el rendimiento de estos últimos. Después de la entrada pasamos a la primera capa oculta, de un total de cuatro que tiene la red. Estas se componen de 32, 64, 128 y 256 neuronas, siguiendo un orden de izquierda a derecha, es decir, de la entrada a la salida. Por último, encontramos una capa de salida con tantas dimensiones como movimientos a clasificar. Generalmente, estos fueron 4: *“Initial Stance”*, *“Medium Stance”*, *“Final Stance”* y *“Swing”*. Sin embargo, como podrá verse a continuación, en algunos de los entrenamientos donde solo se usó como *dataset* los pacientes pediátricos con parálisis cerebral se añadieron dos movimientos a mayores: *“PC Initial Stance”* y *“PC Swing”*. Como función de activación, se hizo uso de la denominada ReLU, excepto en la última capa, donde se empleó la función Softmax. Para evitar situaciones de *overfitting* se aplicó el método de regularización *Dropout* entre las diferentes capas ocultas, con una probabilidad de desactivación igual a 0.2. Además, se usó *Adam* como optimizador, con una tasa de aprendizaje con un valor de 10^{-3} , y *sparse categorical cross entropy* como función de pérdida, usada en tareas de clasificación multiclase. Por último, se contó con los métodos de regularización *Early Stopping* y *Reduce Learning Rate on Plateau*, donde se

seleccionó el error en el grupo de validación como métrica a controlar y un valor de paciencia igual a 3. En la Figura 6.6, se ha tratado de ilustrar de la manera más visual y clara posible esta arquitectura de red.

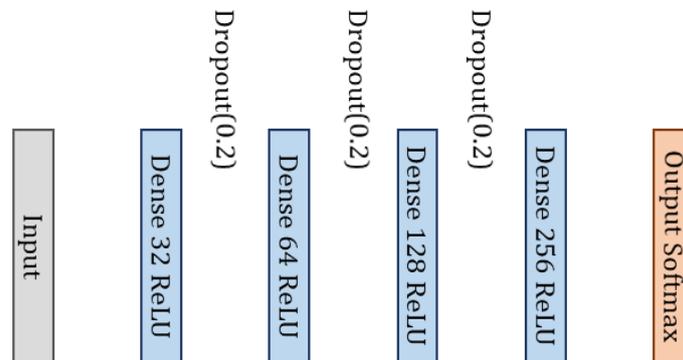


Figura 6.6. Arquitectura DNN desarrollada.

- CNN

El segundo modelo creado se correspondía con una estructura de red convolucional. Este estaba compuesto por 4 bloques principales, cada uno de cuales estaba formado a su vez por la siguiente secuencia de capas: convolucional, *max-pooling* y *Batch Normalization*. En todas las capas convolucionales, cuyo número de neuronas se muestra en la Figura 6.6, se aplicó la estrategia de relleno *padding*, así como se utilizó un parámetro de *stride* igual a uno y *ReLU* como función de activación. Además, el tamaño del *kernel* fue $4 \times 4 \times 1$, excepto en la primera de las cuatro capas, en la cual tenía unas dimensiones de $3 \times 3 \times 1$. Las capas de *max-pooling* se caracterizaron por emplear un tamaño de ventana 2×2 y un *stride* de 2, y por aplicar también la operación de *padding*. Tras estos 4 bloques, se añade el método de regularización *Dropout* con un valor de 0.3 como parámetro. A continuación, se aplica la operación de *Flatten()* para transformar los datos bidimensionales en un único vector unidimensional. Por último, encontramos una capa oculta formada por 516 neuronas y la capa de salida, en las que se utilizó como función de activación la Sigmoide y Softmax, respectivamente. Cabe mencionar que se aplicó el mismo optimizador, función de pérdidas, valor de tasa de aprendizaje y configuración para los métodos de regularización *Early Stopping* y *Reduce Learning Rate on Plateau* que en la red DNN.

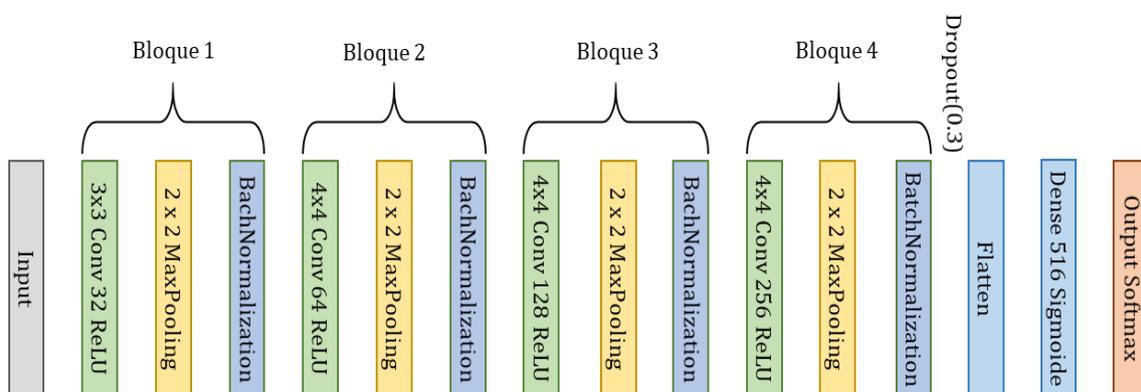


Figura 6.7. Arquitectura CNN desarrollada.

6.4.3 Entrenamientos implementados

Para la realización de este estudio, se dividió el conjunto de entrenamientos implementados en 4 bloques principales. En cada uno de ellos dos clasificadores fueron entrenados, uno para indicar la fase de la marcha en la que se encuentra el pie derecho, y otro para el pie izquierdo.

- **Bloque 1 “Entrenamientos con sujeto adulto sano”.** Tanto el conjunto de test, como el de validación y test se conformó por los voluntarios adultos de control, es decir, los sujetos S01-S10 y S12-S15. La razón de existencia de este primer bloque fue comprobar si una red sencilla con pocos parámetros y variables era capaz de llevar a cabo la tarea de segmentación de marcha.
- **Bloque 2 “Entrenamiento con sujetos adulto y pediátrico sano”.** En este caso, el conjunto de *train* y validación estaba compuesto por los sujetos adultos de control, mientras que para evaluar si la red realmente estaba aprendiendo a clasificar correctamente se empleó como grupo de test a los 3 pacientes pediátricos de control, es decir, S11, S16 y S17. El objetivo aquí fue examinar si la marcha en adultos y niños, cuando ambos no tienen ninguna clase de patología que afecte a su capacidad motora, es similar.
- **Bloque 3 “Entrenamiento con sujeto sano y patológico”.** Este fue el único bloque en el que se empleó la totalidad de instancias del *dataset*. De esta manera, los sujetos sanos, tanto los adultos como los niños, se distribuyeron entre el conjunto de entrenamiento y validación, y los pacientes pediátricos con parálisis cerebral conformaron el conjunto de test. Esta distribución de muestras puede verse más clara en la Tabla 6.6. Básicamente, con este bloque se trataba de averiguar si una red entrenada únicamente con sujetos sanos era suficiente para clasificar la marcha de pacientes con PC eficientemente, o, si, por el contrario, las grandes diferencias entre ambos grupos hacían de esta tarea una imposible.

Train	S01, S02, S03, S04, S05, S06, S07, S08, S09, S11, S14, S15, S17
Validación	S10, S13, S12, S16
Test	P01, P02, P03, P04, P05, P06, P07, P08

Tabla 6.7. Distribución de las muestras en el Bloque 3 de entrenamientos.

- **Bloque 4 “Entrenamiento con sujeto pediátrico con PC”.** Para este último conjunto de entrenamientos solo se emplearon las muestras pertenecientes a pacientes patológicos, tanto para entrenar y validar la red, como para evaluarla. La finalidad fue comprobar si una red entrenada solo con sujetos con PC captaba mejor las particularidades de la marcha en este tipo de pacientes y, por lo tanto, conseguía unos resultados más eficientes que los obtenidos durante el Bloque 3.

Otro de los objetivos perseguidos durante estos entrenamientos fue hallar cuál era el número mínimo de variables necesarias para alcanzar un rendimiento óptimo de la red, es decir, cuáles de ellas aportaban la información más valiosa. Por ello, a lo largo de los distintos bloques de entrenamiento se fueron probando los siguientes subconjuntos de variables:

- **Conjuntos con variables del miembro derecho (R) y miembro izquierdo (L).**

ALLBODY(RL)	<i>"pelvis_tilt", "pelvis_list", "pelvis_rotation", "hip_flexion_r", "hip_adduction_r", "hip_rotation_r", "knee_angle_r", "ankle_angle_r", "subtalar_angle_r", "mtp_angle_r", "hip_flexion_l", "hip_adduction_l", "hip_rotation_l", "knee_angle_l", "ankle_angle_l", "subtalar_angle_l", "mtp_angle_l", "lumbar_extension", "lumbar_bending", "lumbar_rotation", "arm_flex_r", "arm_add_r", "arm_rot_r", "elbow_flex_r", "pro_sup_r", "arm_flex_l", "arm_add_l", "arm_rot_l", "elbow_flex_l", "pro_sup_l"</i>
LOWERBODY(RL)	<i>"pelvis_tilt", "pelvis_list", "pelvis_rotation", "hip_flexion_r", "hip_adduction_r", "hip_rotation_r", "knee_angle_r", "ankle_angle_r", "subtalar_angle_r", "mtp_angle_r", "hip_flexion_l", "hip_adduction_l", "hip_rotation_l", "knee_angle_l", "ankle_angle_l", "subtalar_angle_l", "mtp_angle_l", "lumbar_extension", "lumbar_bending", "lumbar_rotation"</i>
KAH(RL)	<i>"knee_angle_r", "ankle_angle_r", "hip_flexion_r", "knee_angle_l", "ankle_angle_l", "hip_flexion_l"</i>
KA(RL)	<i>"knee_angle_r", "ankle_angle_r", "knee_angle_l", "ankle_angle_l"</i>
KH(RL)	<i>"knee_angle_r", "hip_flexion_r", "knee_angle_l", "hip_flexion_l"</i>

Tabla 6.8. Subconjuntos de variables con información del miembro derecho e izquierdo.

- **Conjuntos con variables solo del miembro derecho (R).**

KA(R)	<i>"knee_angle_r", "ankle_angle_r"</i>
KAH(R)	<i>"knee_angle_r", "ankle_angle_r", "hip_flexion_r"</i>
KAE(R)	<i>"knee_angle_r", "ankle_angle_r", "elbow_flex_r"</i>
KAA(R)	<i>"knee_angle_r", "ankle_angle_r", "arm_flex_r"</i>

Tabla 6.9. Subconjuntos de variables con información solo del miembro derecho.

- **Conjuntos con variables solo del miembro izquierdo (L).**

KA(L)	<i>"knee_angle_l", "ankle_angle_l"</i>
KAH(L)	<i>"knee_angle_l", "ankle_angle_l", "hip_flexion_l"</i>
KAE(L)	<i>"knee_angle_l", "ankle_angle_l", "elbow_flex_l"</i>
KAA(L)	<i>"knee_angle_l", "ankle_angle_l", "arm_flex_l"</i>

Tabla 6.10. Subconjuntos de variables con información solo del miembro izquierdo.

Por último, comentar que, con la intención de mejorar el rendimiento de la red, así como acelerar su aprendizaje, varios hiperparámetros fueron modificados durante los sucesivos entrenamientos. Estos fueron: longitud de la ventana de segmentación (V), número de épocas (E) y *batch size* (BS), es decir, el número de muestras que serán utilizadas a la vez para obtener el error de la función de pérdidas y, en consecuencia, para calcular las actualizaciones de los pesos. Asimismo, para decidir qué configuraciones de entrenamientos funcionaban mejor, se calcularon las métricas de sensibilidad (SEN), especificidad (ESP), *accuracy* (ACC) y *F1-Score* sobre el conjunto de test.

7. RESULTADOS

A continuación, se muestran las distintas configuraciones de entrenamiento implementadas, indicando el subconjunto de variables, arquitectura de red e hiperparámetros seleccionados. A mayores, se presentan los valores obtenidos sobre el conjunto de test en las distintas métricas de rendimiento. Además, comentar que se completó la información sobre los entrenamientos recogida en las tablas, añadiendo los datos finales de *accuracy* alcanzados en el grupo de *train* (ACC-Train) y validación (ACC-Val), junto con la época en la que se detuvo el entrenamiento, en el caso de que hubiese saltado *Early Stopping* (E-Stop).

7.1 Bloque 1. Entrenamientos con sujeto adulto sano

7.1.1 Clasificador Derecho

Nº	Red	Variables	Hiperparámetros			Métricas				Comentarios		
			V	BS	E	SEN	ESP	ACC	F1-SCORE	ACC-Train	ACC-Val	E-STOP
1	DNN	LOWERBODY(RL)	30	4	30	1	1	1	1	1	1	-
2	DNN	LOWERBODY(RL)	30	16	10	1	1	1	1	1	1	-
3	DNN	KA(RL)	30	4	30	0.94	0.98	0.98	0.96	1	1	25
4	DNN	KA(RL)	30	16	10	0.96	0.98	0.98	0.96	1	1	-
5	DNN	KA(R)	30	4	30	0.95	0.98	0.98	0.96	0.99	0.92	18
6	DNN	KA(R)	15	4	30	0.96	0.98	0.98	0.97	0.97	0.85	20
7	DNN	KAE(R)	30	15	4	0.96	0.99	0.99	0.97	0.99	0.83	15
8	DNN	KAA(R)	30	15	4	0.98	0.99	0.99	0.98	1	0.93	12

Tabla 7.1. Entrenamientos del Bloque 1 del clasificador derecho.

Como puede verse recogido en la Tabla 7.1, se inició el entrenamiento usando la red DNN y empleando la información proporcionada por todas las variables correspondientes al tren inferior del cuerpo (1). Con esta sencilla configuración las métricas de rendimiento fueron iguales a la unidad, por lo que se trató de reducir el tiempo de entrenamiento aumentando el tamaño del *batch* y disminuyendo el número de épocas (2). Puesto que, aun cambiando los valores de estos hiperparámetros, los resultados siguieron siendo igual de buenos, la estrategia pasó a ser descubrir el mínimo número de variables necesarias para alcanzar un rendimiento óptimo.

Siguiendo nuestro propio criterio, dado que durante el etiquetado de datos mediante la interfaz de Python la flexión de la rodilla y del tobillo fueron los aspectos usados, principalmente, para determinar el inicio y final de cada fase, se decidió emplear estos ángulos como únicas variables del modelo (3 y 4). Así pues, se logró pasar de una red inicial con 33924 parámetros a una de 18564, es decir, de prácticamente la mitad de ellos.

Incluso con esta gran reducción de variables, las métricas obtenidas siguieron siendo de un valor prácticamente igual a 1. Ante ello, se propuso la hipótesis de si la información acerca del movimiento de la rodilla y tobillo de tan solo el mismo miembro, en este caso de la extremidad derecha, era suficiente para aprender y ejecutar la tarea de clasificación correctamente. De este modo, únicamente se usaron dos variables para entrenar el modelo, la flexión de la rodilla y del tobillo derecho, pasando así a una red más simple de 16664 parámetros (5). Por primera vez, la *accuracy* del conjunto de entrenamiento y validación no fue de 1, sin embargo, era muy cercana a este valor y las métricas obtenidas en el grupo de test seguían siendo de la misma calidad que las alcanzadas en entrenamientos anteriores.

A fin de intentar mejorar estos datos se probaron dos estrategias diferentes. Por una parte, aumentar el número de muestras existentes para entrenar la red, disminuyendo el tamaño de ventana a 15 con un solapamiento del 33%, no obstante, los resultados apenas variaron (6). Por otro lado, se añadió información sobre el movimiento de la extremidad superior derecha, ya que, al menos durante las capturas de sujetos adultos sanos, se contempló como el balanceo de brazos acompañaba al movimiento de marcha de los pies. Así pues, se probó esta teoría con dos nuevas variables: flexión del codo (7) y flexión del brazo derecho (8). Debido a estos cambios, nuestra red pasó a tener un nuevo total de 17604 parámetros ajustables. De entre las dos características, la flexión del brazo fue la que obtuvo unas métricas superiores en todos los grupos, siendo además mejores que aquellas que se obtuvieron únicamente con la información de la flexión de la rodilla y tobillo derecho.

Una vez recogidos todos estos experimentos, en aquellos con mejores resultados (marcados con un color más oscuro en la Tabla 7.1), se procedió a configurar un entrenamiento de tipo *k-fold cross validation*. La razón de todo ello reside en que, dado que nuestra base de datos tiene un tamaño limitado y el grupo de test tan solo se compone de dos sujetos, podría coincidir que justamente estos dos caminasen de una manera similar que aquellos con los que se entrena y que, por ello, se obtengan las buenas métricas presentadas anteriormente. Así pues, para comprobar si los mejores resultados se tratan de una simple casualidad y que ante un conjunto de test diferente la red no es capaz de generalizar igual de bien, aplicamos la estrategia de *k-fold cross validation*. Concretamente, se generaron cinco subconjuntos, cada uno de los estaba caracterizado por la siguiente configuración de sujetos:

Nº	Train	Validación	Test
1	S04, S05, S06, S07, S08, S09, S10, S11, S12, S13, S14, S15	S01	S02, S03
2	S01, S02, S03, S07, S08, S09, S10, S11, S12, S13, S14, S15	S04	S05, S06
3	S01, S02, S03, S04, S05, S06, S10, S11, S12, S13, S14, S15	S07	S08, S09
4	S01, S02, S03, S04, S05, S06, S07, S08, S09, S13, S14, S15	S10	S11, S12
5	S01, S02, S03, S04, S05, S06, S07, S08, S09, S10, S11, S12	S13	S14, S15

Tabla 7.2. Subconjuntos creados para implementar *k-fold cross validation* para el Bloque 1.

Los resultados obtenidos se recogen en la siguiente tabla:

Nº	Red	Variables	Hiperparámetros			Métricas			
			V	BS	E	SEN	ESP	ACC	F1-SCORE
1	DNN	LOWERBODY(RL)	30	4	30	1	1	1	1
3	DNN	KA(RL)	30	4	30	0.98	1	0.99	0.98
5	DNN	KA(R)	30	4	30	0.91	0.97	0.96	0.91
8	DNN	KAA(R)	30	4	30	0.93	0.97	0.96	0.92
9	DNN	KAH(R)	30	4	30	0.98	1	0.99	0.99

Tabla 7.3. Entrenamientos *k-fold cross validation* para el clasificador derecho del Bloque 1.

Analizando los resultados recogidos en la anterior tabla podemos comprobar cómo, en el caso del primer y tercer tipo de entrenamiento, es decir, aquellos que usaban por un lado todas las variables correspondientes a las extremidades inferiores y, por el otro, solo los ángulos de la rodilla y tobillo, el rendimiento no bajó de orden. Más aún, para la tercera configuración de entrenamiento los valores de todas las métricas mejoraron. Así pues, en ambos casos, los anteriores buenos resultados no se debían a una simple coincidencia. Sin embargo, en los otros dos restantes, es decir, aquellos en los que únicamente se utilizó información sobre el propio miembro derecho, a pesar de que las métricas obtenidas fueron buenas, es cierto que sus valores disminuyeron al implementar *k-fold cross validation*.

A fin de tratar de mejorar esos últimos valores se probó con una última combinación de variables. Concretamente, a la flexión de rodilla y tobillo derecho, se les añadieron datos sobre la flexión de cadera de este mismo lado (9). Inspeccionando la Tabla 7.3, se puede comprobar como este nuevo conjunto de variables tiene un gran potencial, ya que, también usando un enfoque de validación cruzada, ha conseguido elevar de nuevo las métricas, y todo ello sin usar información sobre el miembro contralateral, esto es, el izquierdo.

Por último, se quiso examinar como el uso de los mismos ángulos, pero tan solo del miembro contrario, es decir, del izquierdo, afectaba al rendimiento de la red. Sorprendentemente, este no empeoró o se mantuvo igual, sino que mejoró.

Nº	MIEMBRO DERECHO (PROPIO)					MIEMBRO IZQUIERDO (CONTRARIO)				
	Variables	SEN	ESP	ACC	F1-SCORE	Variables	SEN	ESP	ACC	F1-SCORE
5	KA(R)	0.91	0.97	0.96	0.91	KA(L)	0.96	0.99	0.98	0.96
8	KAA(R)	0.93	0.97	0.96	0.92	KAA(L)	0.95	0.99	0.98	0.96
9	KAH(R)	0.98	1	0.99	0.99	KAH(L)	1	1	1	1

Tabla 7.4. Comparación de entrenamientos *k-fold cross validation* para clasificador derecho en Bloque 1 según el lado usado como fuente de información.

7.1.2 Clasificador izquierdo

Nº	Red	Variables	Hiperparámetros			Métricas				Comentarios		
			V	BS	E	SEN	ESP	ACC	F1-SCORE	ACC-Train	ACC-Val	E-STOP
1	DNN	LOWERBODY(RL)	30	4	30	1	1	1	1	1	1	-
2	DNN	LOWERBODY(RL)	30	16	10	1	1	1	1	1	1	-
3	DNN	KA(RL)	30	4	30	0.99	0.99	0.99	0.99	1	1	-
4	DNN	KA(RL)	30	16	10	0.98	0.99	0.99	0.98	1	1	-
5	DNN	KA(L)	30	4	30	0.98	0.99	0.99	0.99	1	0.96	14
6	DNN	KA(L)	15	4	30	0.93	0.98	0.97	0.92	1	0.94	27
7	DNN	KAE(L)	30	4	4	0.95	0.99	0.98	0.95	1	0.97	16
8	DNN	KAA(L)	30	4	4	0.96	0.99	0.98	0.97	1	0.94	13

Tabla 7.5. Entrenamientos del Bloque 1 del clasificador izquierdo.

Para entrenar a la red neuronal encargada de la clasificación del pie izquierdo, se siguió un enfoque muy similar al explicado anteriormente con el clasificador derecho. En este caso, también se comenzó a evaluar el modelo haciendo uso de todas las variables correspondientes a las extremidades inferiores del sujeto (1). Ante el buen comportamiento del mismo, incluso disminuyendo el número de épocas y aumentando el tamaño del *batch* (2), se consideró reducir el número de variables usadas y así, con ello, disminuir el número de parámetros de la red, su complejidad y riesgo de *overfitting*. Dado que, aún con únicamente la información acerca de la flexión de ambas rodillas y tobillos (3 y 4), nuestra DNN obtuvo unas métricas de rendimiento excelentes, se prosiguió con la eliminación de variables. Al igual que se hizo con el clasificador derecho, se comprobó si la información del propio miembro clasificado, en este caso el izquierdo, era suficiente para alcanzar un buen rendimiento. A diferencia de lo que ocurrió con el otro clasificador, las métricas de rendimiento fueron prácticamente iguales a la unidad usando tan solo esas dos características, es decir, la flexión de la rodilla y tobillo izquierdo (5). Es más, incrementando el número de muestras (6) y añadiendo información sobre el movimiento del codo (7) o brazo izquierdo (8), la efectividad de la red decreció, al contrario de lo que sucedió en el clasificador de miembro derecho ante estas últimas dos estrategias.

De nuevo, para comprobar que estos resultados no se trataban de una simple casualidad, se programaron entrenamientos *k-fold cross validation* en aquellos con las mejores métricas (marcados en color más oscuro en la Tabla 7.5). Los subconjuntos de sujetos empleados son los mismos que los anteriormente indicados en la Tabla 7.2. Asimismo, ya que previamente se pudo ver como los datos sobre la flexión de la cadera aportaban valiosa información al

modelo, se experimentó directamente con el conjunto de variables denominados como KAH(L).

Nº	Red	Variables	Hiperparámetros			Métricas			
			V	BS	E	SEN	ESP	ACC	F1-SCORE
1	DNN	LOWERBODY(RL)	30	4	30	1	1	1	1
3	DNN	KA(RL)	30	4	30	0.99	1	1	0.99
5	DNN	KA(L)	30	4	30	0.96	0.99	0.98	0.96
9	DNN	KAH(L)	30	4	30	1	1	1	1

Tabla 7.6. Entrenamientos *k-fold cross validation* para el clasificador izquierdo del Bloque 1.

Al igual que como se pudo constatar con el clasificador del miembro derecho, los altos rendimientos registrados para la primera y tercera configuración de hiperparámetros y variables no se debían a las características de los dos sujetos de test elegidos al azar, sino que, realmente, son potentes tipos de entrenamientos con buenas métricas, independientemente de quién conforme el conjunto de test. No obstante, también se ha podido comprobar como esta última afirmación no se cumple para el quinto tipo de entrenamiento, en el cual las métricas son menores a las registradas anteriormente. Aun así, hay que destacar que es posible obtener unos valores excelentes con información de únicamente la propia pierna izquierda, gracias a la combinación de flexión de cadera, rodilla y tobillo.

Para concluir esta apartado, se comprobó cómo afecta al rendimiento del clasificador usar el mismo conjunto de variables, pero con información del miembro contrario, es decir, el derecho. Al contrario de lo que sucedió en el clasificador derecho, en este caso, se obtienen mejores métricas cuando la información es proporcionada por el propio miembro que está siendo clasificado, esto es, el izquierdo.

Nº	MIEMBRO IZQUIERDO (PROPIO)					MIEMBRO DERECHO (CONTRARIO)				
	Variabes	SEN	ESP	ACC	F1-SCORE	Variabes	SEN	ESP	ACC	F1-SCORE
5	KA(L)	0.96	0.99	0.98	0.96	KA(R)	0.90	0.97	0.95	0.90
8	KAH(L)	1	1	1	1	KAH(R)	0.97	0.99	0.99	0.97

Tabla 7.7. Comparación de entrenamientos *k-fold cross validation* para clasificador izquierdo en Bloque 1 según el lado usado como fuente de información.

7.2 Bloque 2. Entrenamientos con sujeto adulto y pediátrico sano

7.2.1 Clasificador derecho

Nº	Red	Variables	Hiperparámetros			Métricas				Comentarios		
			V	BS	E	SEN	ESP	ACC	F1-SCORE	ACC-Train	ACC-Val	E-STOP
1	DNN	ALLBODY(RL)	30	4	30	0.93	0.97	0.97	0.93	1	1	-
2	DNN	LOWERBODY(RL)	30	4	30	0.95	0.99	0.98	0.96	1	1	-
3	DNN	KA(RL)	30	4	30	0.87	0.96	0.93	0.84	1	0.99	16
4	DNN	KA(R)	30	4	30	0.78	0.93	0.87	0.76	1	0.95	13
5	DNN	KAA(R)	30	4	30	0.80	0.94	0.91	0.78	1	0.92	17
6	DNN	KAH(R)	30	4	30	0.90	0.97	0.96	0.91	1	0.97	16
7	CNN	KA(RL)	30	4	30	1	1	1	1	1	1	-
8	CNN	KA(RL)	30	16	12	0.99	0.99	0.99	0.99	1	1	-
9	CNN	KA(R)	30	4	30	0.87	0.97	0.95	0.88	1	1	29
10	CNN	KAH(R)	30	4	30	0.93	0.98	0.98	0.94	1	1	20

Tabla 7.8. Entrenamientos del Bloque 2 del clasificador derecho.

En un primer momento, dado que se partió de la idea de que la marcha de un adulto y de un niño, ambos sanos, guardaban diferencias entre sí, se entrenó una red utilizando toda la información disponible, es decir, haciendo uso de las 30 variables existentes (1). Debido a ello, la DNN, aun siendo una red muy simple con tan solo 4 capas ocultas, tenía un gran número de parámetros, exactamente 43524. Los resultados para el grupo de test fueron buenos, sin embargo, eran mejorables, por lo que se probó a eliminar aquellas variables relacionadas con movimientos del tren superior (2). La razón de la anterior iniciativa se fundamenta en que, al examinar con detenimiento los videos de la marcha de los sujetos pediátricos, se comprobó que el braceo del S16 no era sincrónico ni fluido, como lo era el de las capturas en adultos. Seguramente, esto fuese debido al propio ambiente de grabación, donde los niños han podido sentirse más tensos y vigilados que cuando andan normalmente por la calle o por sus casas.

Tras esta eliminación de variables, el rendimiento incrementó. Ante este éxito, se intentó reducir aún más el número de características, limitando estas a las 4 recogidas bajo el nombre de KA(ALL) (3). Por desgracia el resultado no fue tan bueno como lo esperado. Asimismo, tampoco lo fue ante el uso de únicamente la información proporcionada por la flexión de rodilla y tobillo del miembro derecho (4). Añadir como variable la flexión del brazo derecho (5) apenas supuso mejoría, tal y como era de esperar por lo explicado en el anterior párrafo sobre el movimiento de brazos del S16. Lo que sí que logró incrementar los valores de las métricas fue emplear, a mayores, información sobre la flexión de la cadera derecha (6). Aun así, no nos conformamos con estos últimos resultados, por lo que la

estrategia paso a ser probar otro tipo de estructura de red diferente, la red CNN explicada con anterioridad.

Gracias a esta red nueva red, mucho más compleja que una simple DNN, el valor alcanzado en todas las métricas de rendimiento sobre el conjunto de test, usando como variables el subconjunto KA(RL) (7), fue igual a la unidad y, de prácticamente el mismo orden, incluso incrementando el tamaño del batch y reduciendo el número de épocas a tan solo 12 (8). En vista de los resultados, se volvió a tratar de disminuir el número de variables empleadas. Tan solo usando la información acerca de la flexión de la rodilla y tobillo derecho (9), la efectividad de la red decreció, especialmente los valores de las métricas de especificidad y *accuracy*. No obstante, estas volvieron a aumentar de nuevo al incorporar datos sobre el movimiento de flexión de la cadera derecha, es decir, empleando el subconjunto de variables KAH(R) (10).

En este caso, dado que únicamente se pudo registrar a tres sujetos pediátricos sanos (S11, S16 y S17), no fue posible implementar entrenamientos con validación cruzada, por lo que se pudo comprobar si los buenos resultados se deben a la coincidencia de que el tipo de marcha de estos tres niños es similar al de los adultos con los que se entrenó la red, o si, por el contrario, realmente hemos desarrollado una buena configuración de red. Lo que sí que hicimos, continuando con un enfoque similar al seguido en el anterior bloque de entrenamiento, fue comparar las métricas obtenidas en función de si las variables empleadas se correspondían con solo movimientos de la pierna derecha o izquierda. De nuevo, tal y como puede comprobarse en la Tabla 7.9, se obtuvieron mejores resultados en el clasificador derecho a partir de la información proporcionada por variables del miembro izquierdo.

Nº	MIEMBRO DERECHO (PROPIO)					MIEMBRO IZQUIERDO (CONTRARIO)				
	Variables	SEN	ESP	ACC	F1-SCORE	Variables	SEN	ESP	ACC	F1-SCORE
9	KA(R)	0.87	0.97	0.95	0.88	KA(L)	0.90	0.97	0.95	0.90
10	KAH(R)	0.93	0.98	0.98	0.94	KAH(L)	0.99	1	0.99	0.98

Tabla 7.9. Comparación de entrenamientos para clasificador derecho en Bloque 2 según el lado usado como fuente de información.

7.2.2 Clasificador izquierdo

Nº	Red	Variables	Hiperparámetros			Métricas				Comentarios		
			V	BS	E	SEN	ESP	ACC	F1-SCORE	ACC-Train	ACC-Val	E-STOP
1	DNN	ALLBODY(RL)	30	4	30	0.94	0.98	0.98	0.94	1	1	-
2	DNN	LOWERBODY(RL)	30	4	30	0.82	0.94	0.91	0.80	1	1	-
3	DNN	KA(RL)	30	4	30	0.80	0.94	0.91	0.80	1	0.99	22
4	DNN	KA(L)	30	4	30	0.58	0.86	0.78	0.60	1	0.95	16
5	DNN	KAA(L)	30	4	30	0.70	0.88	0.81	0.64	1	0.96	16
6	DNN	KAH(L)	30	4	30	0.90	0.97	0.96	0.90	1	0.97	15
7	CNN	KA(RL)	30	4	30	0.99	1	1	0.99	1	1	-
8	CNN	KA(RL)	30	16	12	0.91	0.97	0.95	0.90	1	1	-
9	CNN	KA(L)	30	4	30	0.91	0.97	0.96	0.91	1	0.98	14
10	CNN	KAH(L)	30	4	30	0.99	1	0.99	0.98	1	1	-

Tabla 7.10. Entrenamientos del Bloque 2 del clasificador izquierdo.

Para el caso del clasificador del pie izquierdo, las métricas sobre el conjunto de test obtuvieron mejores valores haciendo uso de las 30 variables (1), que utilizando únicamente las relacionadas con las extremidades inferiores (2) o tan solo las de la rodilla y tobillo (3), sin embargo, en los tres casos los resultados eran peores que los obtenidos en el clasificador derecho. Para llevar un orden similar al seguido en el otro clasificador y así poder comparar resultados, se probó una última vez la arquitectura de red DNN, pero empleando tan solo ángulos del lado izquierdo del cuerpo. Se obtuvo un mejor rendimiento con el conjunto de variables KAA(L) (5), que con KA(L) (4), de todos modos, no podrían considerarse entrenamientos exitosos. Lo que sí que logró alcanzar de nuevo unas métricas satisfactorias fue añadir a la flexión de rodilla y tobillo izquierdo, información a mayores sobre la flexión de cadera de esta misma extremidad (6). A pesar de estos últimos hallazgos, se creyó que un cambio en la arquitectura de red aumentaría aún más la efectividad del clasificador.

Ante ello, se probó la red CNN. El rendimiento, usando solo como variables las del conjunto KA(RL), pasó de nuevo a ser excelente, prácticamente igual a la unidad en todas las métricas (7 y 8). Incluso, aun empleando solo información sobre la flexión de la rodilla y tobillo izquierdo (9), los resultados fueron superiores a los obtenidos con la red DNN y un conjunto mayor de variables. Por último, se volvió a comprobar como la flexión de la cadera podría considerarse una variable fundamental, ya que logra incrementar los valores en los resultados de test a unos sobresalientes, y todo ello, tan solo empleando un total de tres variables del miembro izquierdo (10).

Para concluir con esta parte, se comprobó con la información procedente de la pierna izquierda se conseguía un mayor aprendizaje de la red que cuando las variables se correspondían únicamente con datos del lado derecho.

Nº	MIEMBRO IZQUIERDO (PROPIO)					MIEMBRO DERECHO (CONTRARIO)				
	VARIABLES	SEN	ESP	ACC	F1-SCORE	VARIABLES	SEN	ESP	ACC	F1-SCORE
9	KA(L)	0.91	0.97	0.96	0.91	KA(R)	0.94	0.98	0.98	0.94
10	KAH(L)	0.99	1	0.99	0.98	KAH(R)	0.96	0.99	0.98	0.96

Tabla 7.11. Comparación de entrenamientos para clasificador izquierdo en Bloque 2 según el lado usado como fuente de información.

7.3 Bloque 3. Entrenamiento con sujeto sano y patológico

7.3.1 Clasificador derecho

Nº	Red	Variables	Hiperparámetros			Métricas				Comentarios		
			V	BS	E	SEN	ESP	ACC	F1-SCORE	ACC-Train	ACC-Val	E-STOP
1	DNN	ALLBODY(RL)	30	4	30	0.30	0.77	0.65	0.29	1	0.97	23
2	DNN	LOWERBODY(RL)	30	4	30	0.90	0.97	0.96	0.91	1	0.99	15
3	CNN	LOWERBODY(RL)	30	4	30	0.96	0.99	0.98	0.95	1	1	20
4	CNN	LOWERBODY(RL)	30	16	15	0.95	0.98	0.97	0.94	1	1	14
5	CNN	KA(RL)	30	4	30	0.98	0.99	0.99	0.97	1	0.99	22
6	CNN	KA(R)	30	4	30	0.93	0.98	0.97	0.93	1	0.94	15
7	CNN	KAH(R)	30	4	30	0.91	0.97	0.96	0.91	1	0.92	13

Tabla 7.12. Entrenamientos del Bloque 3 del clasificador derecho.

Se inició este nuevo enfoque entrenando la red DNN con todas las variables disponibles (1), sin embargo, como era de esperar, dado que en la mayoría de sujetos pediátricos con parálisis cerebral el movimiento de los brazos fue caótico y no seguía ningún tipo de patrón, el rendimiento fue muy malo, el peor obtenido hasta entonces. Así pues, como respuesta a estos no deseables resultados, se limitó la información aportada a aquellos ángulos correspondientes al tren inferior del cuerpo (2), pasando de una red de 43524 parámetros a otra de 33924. Ya solo con esta simple acción, todas las métricas volvieron a ser iguales o superiores al valor 0.90.

Cambiando la arquitectura de red a la CNN (3) y, como consecuencia, ampliando el número de parámetros a 166520, la efectividad se acrecentó, aun aumentando el tamaño de *batch* y reduciendo el número de épocas (4). Dado que, de nuevo, se alcanzaron unos buenos resultados, se trató de descubrir el número mínimo de variables necesarias para lograr que el rendimiento se mantuviese dentro de ese mismo rango de valores elevados.

Con el uso del subconjunto de variables KA(RL) (5) los valores de las métricas de rendimiento aumentaron, para, de nuevo disminuir unas centésimas, al limitar aún más la información a aquella aportada únicamente por el miembro derecho (6). Además, es vital comentar como, por primera vez, el añadir información sobre la flexión de cadera del

miembro a clasificar, esto es, el derecho (7), no mejoró los resultados, sino que los empeoró. Así pues, podríamos decir que existen diferencias entre el movimiento de cadera de sujetos patológicos y sanos, mucho más notables que las que puedan existir entre los movimientos de rodilla y tobillo, al menos para el lado derecho y para esta base de datos.

En último lugar, comparando los resultados recogidos en la Tabla 7.13, observamos cómo esta fue la primera ocasión en la que las métricas de test del clasificador derecho fueron menores usando solo información del miembro contrario, ya que, hasta este momento, siempre habían sido superiores usando datos sobre el movimiento de la pierna izquierda que sobre la derecha. La causa de esto es el deterioro de la marcha que presentan los pacientes con PC. Concretamente, todos los sujetos grabados presentaban hemiparesia y, en la mayoría, el lado afectado era el izquierdo.

Nº	MIEMBRO DERECHO (PROPIO)					MIEMBRO IZQUIERDO (CONTRARIO)				
	Variables	SEN	ESP	ACC	F1-SCORE	Variables	SEN	ESP	ACC	F1-SCORE
9	KA(R)	0.93	0.98	0.97	0.93	KA(L)	0.88	0.97	0.96	0.90
10	KAH(R)	0.91	0.97	0.96	0.91	KAH(L)	0.88	0.96	0.95	0.88

Tabla 7.13. Comparación de entrenamientos para clasificador derecho en Bloque 3 según el lado usado como fuente de información.

7.3.2 Clasificador izquierdo

Nº	Red	Variables	Hiperparámetros			Métricas				Comentarios		
			V	BS	E	SEN	ESP	ACC	F1-SCORE	ACC-Train	ACC-Val	E-Stop
1	DNN	ALLBODY(RL)	30	4	30	0.27	0.76	0.67	0.27	1	0.90	11
2	DNN	LOWERBODY(RL)	30	4	30	0.61	0.86	0.81	0.61	1	0.98	14
3	CNN	LOWERBODY(RL)	30	4	30	0.65	0.88	0.83	0.67	1	1	-
4	CNN	KA(RL)	30	4	30	0.67	0.88	0.83	0.66	1	1	16
5	CNN	KA(L)	30	4	30	0.68	0.90	0.86	0.70	1	0.94	18
6	CNN	KAH(L)	30	4	30	0.67	0.89	0.85	0.67	1	0.99	15
7	CNN	KH(RL)	30	4	30	0.67	0.89	0.84	0.68	1	1	19

Tabla 7.14. Entrenamientos del Bloque 3 del clasificador izquierdo.

Como ya era de esperar, dado lo comentado con anterioridad acerca de la hemiparesia y afectación del lado izquierdo de los sujetos pediátricos patológicos, el rendimiento de este clasificador fue inferior al del lado derecho. Los peores resultados se alcanzaron con la red DNN al hacer uso de todo el conjunto de variables disponibles (1), debido a la falta de coordinación entre el movimiento de piernas y balanceo de brazos en los sujetos con PC. La efectividad mejoró al considerar solo los ángulos de los miembros inferiores del cuerpo (2), sin embargo, aun así, esta era baja. Lo más sorprendente fue la escasa mejoría que experimentó el clasificador al cambiar a la arquitectura de red CNN (3 y 4), ya que, en el

resto de pruebas realizadas previamente siempre había supuesto una mejora notable. De nuevo, tampoco datos acerca del ángulo de flexión de la cadera (6) o eliminar la variable de flexión del tobillo (7) consiguieron aumentar el valor de las métricas. Así pues, cuando el conjunto de test lo conforman sujetos pediátricos con la condición de parálisis cerebral, ha podido verse como los mejores resultados se consiguen cuando se usa el menor número de variables, en este caso, flexión de la rodilla y tobillo izquierdo (5), de todos modos, estos se encuentran muy lejos de poder considerarse como buenos.

Conforme a lo esperado, las métricas presentaron valores más elevados al utilizarse la información del miembro derecho, lado no afectado en la mayoría de los niños y, por lo tanto, con movimientos más similares al de los sujetos sanos. Aun así, los resultados de este último clasificador no fueron de nuestro conformismo, por lo que, a continuación, en el siguiente bloque, se probará a entrenar un clasificador únicamente con sujetos patológicos, a fin de intentar mejorar los obtenidos para este último.

Nº	MIEMBRO IZQUIERDO (PROPIO)					MIEMBRO DERECHO (CONTRARIO)				
	Variables	SEN	ESP	ACC	F1-SCORE	Variables	SEN	ESP	ACC	F1-SCORE
9	KA(L)	0.68	0.90	0.86	0.70	KA(R)	0.69	0.90	0.87	0.70
10	KAH(L)	0.67	0.89	0.85	0.67	KAH(R)	0.71	0.91	0.88	0.72

Tabla 7.15. Comparación de entrenamientos para clasificador izquierdo en Bloque 3 según el lado usado como fuente de información.

7.4 Bloque 4. Entrenamiento con sujeto pediátrico con PC

7.4.1 Clasificador derecho

Nº	Red	Variables	Hiperparámetros			Métricas				Comentarios		
			V	BS	E	SEN	ESP	ACC	F1-SCORE	ACC-Train	ACC-Val	E-STOP
1	DNN	ALLBODY(RL)	30	4	30	0.98	1	0.99	0.99	1	0.93	13
2	DNN	LOWERBODY(RL)	30	4	30	0.99	1	1	0.99	1	1	-
3	DNN	KA(RL)	30	4	30	0.87	0.97	0.95	0.87	1	1	-
4	CNN	LOWERBODY(RL)	30	4	30	1	1	1	1	1	1	-
5	CNN	KA(RL)	30	4	30	0.96	0.99	0.99	0.96	1	1	22
6	CNN	KAH(R)	30	4	30	1	1	1	0.99	1	0.99	16
7	CNN	KA(R)	30	4	30	0.93	0.98	0.97	0.93	1	1	16

Tabla 7.16. Entrenamientos del Bloque 4 del clasificador derecho.

En este nuevo bloque de entrenamientos, donde la base de datos utilizada se redujo a la de pacientes pediátricos patológicos, se puede comprobar como los resultados obtenidos son mucho mayores que cuando se empleaban, a mayores, los sujetos sanos, tanto adultos como niños. De esta manera, en la configuración de entrenamiento en la que se usaron todas las

variables (1) vemos como las métricas son excelentes, contrastando con los casos anteriores en los que también se había empleado información sobre el tren superior del cuerpo. Esto es así porque, prácticamente, todos los sujetos de la base de datos con parálisis cerebral se caracterizan por tener un movimiento de brazos caótico, el cual, a su vez, es muy distinto al que tienen los individuos de control. A pesar de los buenos resultados obtenidos con este subconjunto ALLBODY(RL), se hicieron más pruebas a fin de reducir el número de variables empleadas.

Mientras que con LOWERBODY(RL) (2) el rendimiento de la DNN seguía siendo óptimo, con tan solo las cuatro variables incluidas en KA(RL) (3) este descendió notablemente. Por ello, tal y como se hizo en ocasiones previas, se sustituyó la red utilizada por la CNN. Con esta nueva arquitectura, los resultados volvieron a ser de nuevo óptimos. Destacar que, tan solo haciendo uso de la flexión de rodilla, tobillo y cadera de la extremidad derecha (6), se obtuvieron unas métricas igual de buenas que las adquiridas empleando todas las variables correspondientes a movimientos del tren inferior del cuerpo (4). Con esta red CNN, el rendimiento más bajo se alcanzó con el subconjunto KA(R) (7), no obstante, en líneas generales, este puede considerarse como bueno.

Tal y como se hizo cuando solo se empleó como base de datos el conjunto de adultos sanos, esto es, en el Bloque 1, se repitieron los entrenamientos con mayor interés (marcados con un color más oscuro en la Tabla 7.16) haciendo validación cruzada, con el propósito de comprobar que los buenos resultados que se alcanzaron no eran una simple casualidad, fruto de la disposición de los sujetos P01-P08 en los conjuntos de entrenamiento, validación y test. A su vez, se aprovechó para comparar el efecto que causaba en el clasificador el hacer solo uso de variables del miembro derecho o del miembro izquierdo. A continuación, se muestran los cuatro conjuntos de instancias generados:

Conjunto	Train	Validación	Test
1	P04, P05, P06, P07, P08	P03	P01, P02
2	P01, P02, P06, P07, P08	P05	P03, P04
3	P01, P02, P03, P04, P08	P07	P05, P06
4	P02, P03, P04, P05, P06	P01	P07, P08

Tabla 7.17. Subconjuntos creados para implementar *k-fold cross validation* para el Bloque 4.

Nº	Red	Variables	Hiperparámetros			Métricas			
			V	BS	E	SEN	ESP	ACC	F1-SCORE
5	CNN	KA(RL)	30	4	30	0.91	0.98	0.97	0.91
ENTRENAMIENTOS CON VARIABLES DE SOLO EL MIEMBRO DERECHO (PROPIO)									
6	CNN	KAH(R)	30	4	30	0.92	0.97	0.96	0.91
7	CNN	KA(R)	30	4	30	0.77	0.94	0.91	0.76
ENTRENAMIENTOS CON VARIABLES DE SOLO EL MIEMBRO IZQUIERDO (CONTRARIO)									
6	CNN	KAH(L)	30	4	30	0.93	0.98	0.97	0.93
7	CNN	KA(L)	30	4	30	0.77	0.93	0.92	0.75

Tabla 7.18. Comparación de entrenamientos *k-fold cross validation* para clasificador derecho en Bloque 4 según el lado usado como fuente de información.

De nuevo, tal y como sucedió en todos los clasificadores derechos, excepto en el del Bloque 3, con los movimientos del lado izquierdo del cuerpo parece que se logra conseguir un mejor aprendizaje de la red.

7.4.2 Clasificador izquierdo

Nº	Red	Variables	Hiperparámetros			Métricas				Comentarios		
			V	BS	E	SEN	ESP	ACC	F1-SCORE	ACC-Train	ACC-Val	E-STOP
1	DNN	ALLBODY(RL)	30	4	30	0.01	0.64	0.50	0.01	0.94	0.81	17
2	DNN	LOWERBODY(RL)	30	4	30	0.06	0.68	0.54	0.05	0.94	0.83	15
3	CNN	LOWERBODY(RL)	30	4	30	0.07	0.66	0.53	0.06	1	0.93	26
4	CNN	KA(RL)	30	4	30	0.1	0.68	0.55	0.09	1	0.93	26
5	CNN	KH(RL)	30	4	30	0.01	0.59	0.47	0.01	1	1	-
6	CNN	KAH(R)	30	4	30	0	0.63	0.5	0	1	1	15

Tabla 7.19. Entrenamientos del Bloque 4 del clasificador izquierdo.

A causa de la afectación en el lado izquierdo en muchos de estos sujetos patológicos, los rendimientos para el clasificador izquierdo fueron muy inferiores a los obtenidos para el caso derecho. Como puede verse en la Tabla 7.19, ni aun cambiando la arquitectura de red a una CNN, la cual tiene mayor complejidad que la DNN y por lo tanto mayor capacidad para extraer características y patrones de los datos, métricas como la sensibilidad y F1-Score no lograron superar el valor de 0.1. Dado que, principalmente la diferencia en la marcha de la pierna izquierda radica en el pie en punta, se probó a eliminar la variable de flexión de tobillo del conjunto (5), sin embargo, ni con esos cambios, se consiguió mejorar la efectividad del clasificador. Tampoco haciendo uso de solo la información del miembro derecho (6), es decir, del lado no afectado, se alcanzó ningún progreso.

Ante todos estos intentos fallidos se decidió apostar por una nueva estrategia. Esta consistía en usar las seis etiquetas, presentadas en el Capítulo 6. Materiales y Métodos, en vez de solo las 4 correspondientes a un patrón de marcha sano. De esta manera, se logró diferenciar cuando la fase de swing y apoyo inicial eran anómalas (etiqueta 5 y 6, respectivamente) y cuando normales (etiquetas 4 y 1, respectivamente). Únicamente los sujetos P02, P05, P07 y P08 tenían la etiqueta 5-6 en alguno de sus *frames*, por lo que los dos primeros (P02 y P05) se incluyeron en el conjunto de *train*, P07 para la validación y P08 en el grupo de test. Así pues, de forma más detallada, los conjuntos formados fueron:

Train	P01, P02, P03, P04, P05
Validación	P07
Test	P06, P08

Tabla 7.20. Distribución de las muestras en el Bloque 4 usando etiquetas 1-6 del clasificador izquierdo.

Los resultados obtenidos se muestran a continuación:

Nº	Red	Variables	Hiperparámetros			Métricas				Comentarios		
			V	BS	E	SEN	ESP	ACC	F1-SCORE	ACC-Train	ACC-Val	E-STOP
1	CNN	LOWERBODY(RL)	30	4	30	0.57	0.94	0.91	0.54	0.99	0.51	12
2	CNN	KAH(RL)	30	4	30	0.63	0.95	0.92	0.56	0.99	0.47	13
3	CNN	KA(RL)	30	4	30	0.68	0.95	0.92	0.65	0.99	0.48	13
4	CNN	KA(L)	30	4	30	0.63	0.95	0.92	0.55	0.99	0.5	16

Tabla 7.21. Entrenamientos del Bloque 4 usando etiquetas 1-6 del clasificador izquierdo.

Con esta acción, el rendimiento mejoró notablemente, sin embargo, cabe destacar como, independientemente del conjunto de variables seleccionadas, las métricas se mantienen prácticamente siempre en torno al mismo valor. Aun así, cabe destacar que los mejores resultados son los obtenidos con el conjunto KA(RL). Para este nuevo caso, no se probó a usar solamente movimientos del lado derecho del cuerpo dado que, únicamente con estos, sería imposible que el clasificador fuese capaz de distinguir cuando en el lado izquierdo se está dando un apoyo o un swing fuera de la normalidad y cuando no.

8. DISCUSIÓN Y LIMITACIONES

En este TFG se han desarrollado y entrenado modelos de *deep learning* con el fin de segmentar la marcha en sus distintas fases. Durante este capítulo, se analizarán las principales ideas extraídas tras ejecutar los 4 bloques de entrenamientos, mostrados en el Capítulo 7. Resultados. Asimismo, se compararán los resultados obtenidos en el presente estudio con los de los otros trabajos presentados previamente en el Capítulo 6. Materiales y Métodos. Para conseguir exponer las ideas con una mayor claridad, se ha decidido dividir la discusión en dos apartados, analizando por un lado la marcha de sujetos sanos, y, por otra parte, la de pacientes con parálisis cerebral. Además, por último, se comentarán brevemente aquellos aspectos que han limitado en parte la ejecución de este TFG.

8.1 Análisis de resultados en marcha de sujeto sano

En primer lugar, centrándonos en el análisis y segmentación de la marcha de personas adultas sanas, podría decirse que el modelo de red DNN desarrollado fue capaz de conseguir unos óptimos resultados, incluso cuando el número de variables era muy reducido. Así pues, teniendo en cuenta los valores en las métricas de los distintos tipos de entrenamiento, se podría afirmar que, con tan solo la información de la flexión de cadera, rodilla y tobillo de uno de los miembros, o bien con la flexión de rodilla y tobillo tanto del lado izquierdo como derecho, bastaría para clasificar correctamente las fases de la marcha en este tipo de sujetos. Muestra de ello, son las siguientes matrices de confusión, obtenidas durante los correspondientes entrenamientos del Bloque 1, en los que se aplicó validación cruzada:

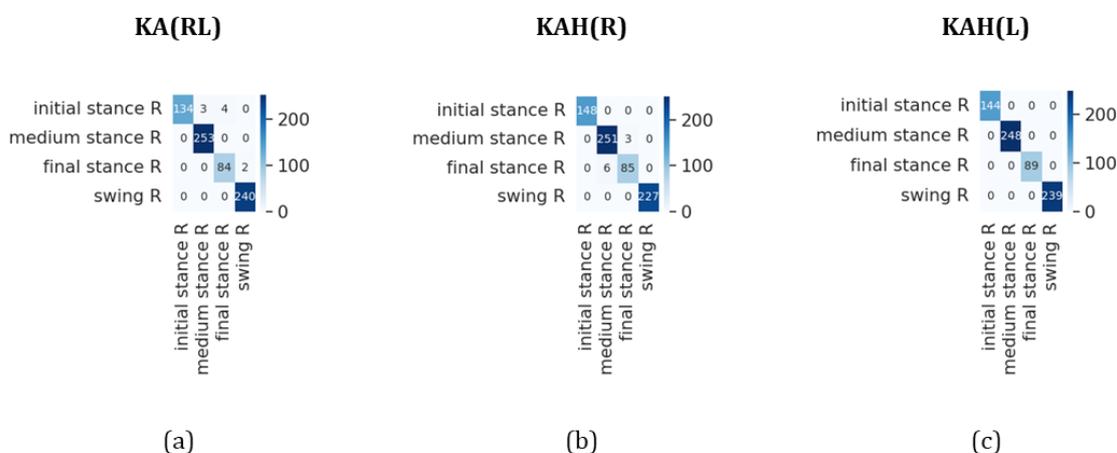


Figura 8.1. Matrices de confusión del Bloque 1 del clasificador derecho. Conjuntos de variables usados: (a) KA(RL), (b) KAH(R), (c) KAH(L).

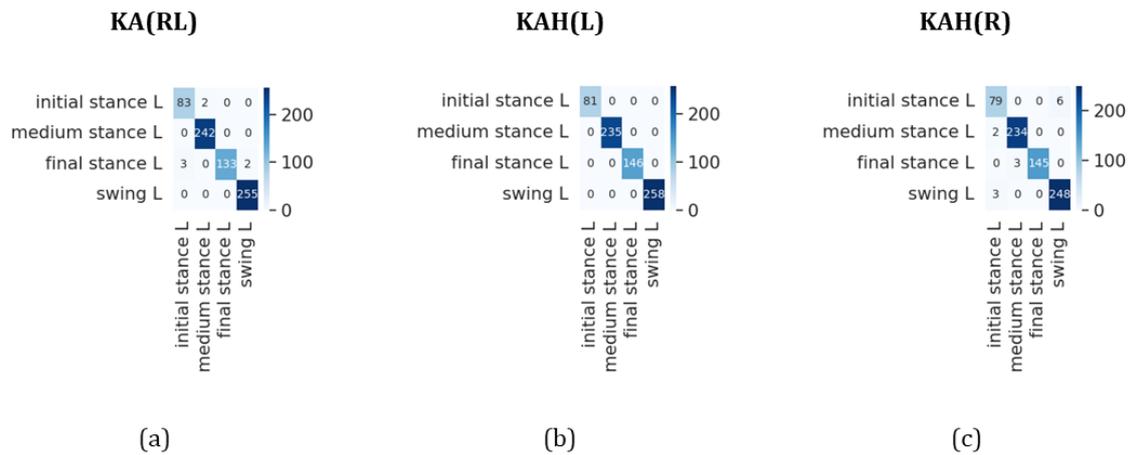


Figura 8.2. Matrices de confusión del Bloque 1 del clasificador izquierdo. Conjuntos de variables usados: (a) KA(RL), (b) KAH(L), (c) KAH(R).

Como puede verse, apenas hay muestras mal clasificadas y, aquellas que fueron predichas erróneamente, fueron confundidas con las fases contiguas. Esto es así porque, el punto que marca el final de una fase y el inicio de la siguiente es un instante que dura milésimas de segundo. En este TFG en concreto, se está abordando un entrenamiento de tipo supervisado en el que las muestras han sido etiquetadas de forma manual. Así pues, a pesar de haber tratado de que este proceso de etiquetado fuese lo más objetivo posible, este se ha hecho inspeccionando de forma visual capturas de vídeo. Por ello, dado que un simple vídeo no proporciona información sobre el valor de ángulo que caracteriza a las articulaciones en cada instante, es muy complicado que en todos los sujetos el punto marcado como final de una fase e inicio de la siguiente siempre se corresponda con los mismos valores, en nuestro caso, de ángulo de flexión, rotación o abducción. Es por ello, que es posible afirmar que la red tiene un funcionamiento óptimo, y que el problema principal se haya en la precisión de etiquetado.

Tras la anterior justificación, se podría concluir con que, con tan solo una simple de red fully-connected de 4 capas ocultas y tres o cuatro variables correspondientes a ángulos de flexión del tren inferior del cuerpo, es posible realizar una tarea de segmentación de la marcha eficientemente. Como consecuencia, la complejidad y número de parámetros de la red disminuye enormemente, siendo la cantidad de estos últimos igual a 18564 o 17604, según se use el conjunto KA(RL) o KAH(R)-KAH(L), respectivamente.

Esto contrasta enormemente con las complejas arquitecturas utilizadas en otros trabajos relacionados, como la red FCN-BiLSTM de Rifaat et al. (2022) o la CNN-GRU de Martindale et al. (2021). En algunos de estos estudios, como para la Gait-Net desarrollada por Song et al. (2019), utilizaban como datos de entrada imágenes, en vez de ángulos articulares, lo cual puede ser la causa de que se requiriesen redes con un mayor número de parámetros entrenables. No obstante, en vista de los resultados obtenidos en el Bloque 1, consideramos mucho más útil y representativa para el movimiento de marcha, aportar como dato a la red los parámetros recogidos por sensores IMU. Más concretamente, nos decantamos por los ángulos articulares. Esto es así porque, en los otros trabajos donde se emplearon dispositivos inerciales, utilizaron como entrada datos de 6 canales, esto es, los 3 ejes del acelerómetro y los 3 ejes del giroscopio (Borzì et al., 2023; Kreuzer & Munz, 2021; Martindale et al., 2021; Rifaat et al., 2022). Ante ello, la dimensionalidad es prácticamente igual de elevada que si se utilizaran las 35 variables correspondientes a nuestro conjunto ALLBODY(RL), aumentando en consecuencia las posibilidades de *overfitting*, enlenteciendo

el proceso de aprendizaje y forzando al uso de redes neuronales complejas y de gran profundidad.

Asimismo, dados los conjuntos que se han considerado, en este TFG, como los necesarios para emplear el mínimo número de variables y, al mismo tiempo, obtener unos óptimos resultados, puede verse como la cantidad de sensores inerciales requeridos se reduce. Utilizando datos de vídeo hemos comprobado como en sujeto adulto sano es posible hacer una correcta segmentación de las fases de la marcha empleando un reducido número de ángulos articulares. Esta información angular puede provenir de ambos lados del cuerpo o de únicamente una de las extremidades inferiores. Este resultado da pie a pensar que es posible realizar un estudio de la marcha empleando un número limitado de sensores. De esta manera, se lograría reducir el tiempo empleado en la colocación de los mismos, así como en la captura de datos. Por consiguiente, nuestra hipótesis sería que, considerando los conjuntos de variables KA(RL), KAH(R) y KAH(L) como los necesarios, únicamente sería preciso usar un total de 4 o 6 sensores. En consecuencia, se reduce en gran medida el número inicial de 17 IMUs utilizadas para la recopilación de nuestra base de datos. Esto puede verse con mayor claridad en la siguiente figura:

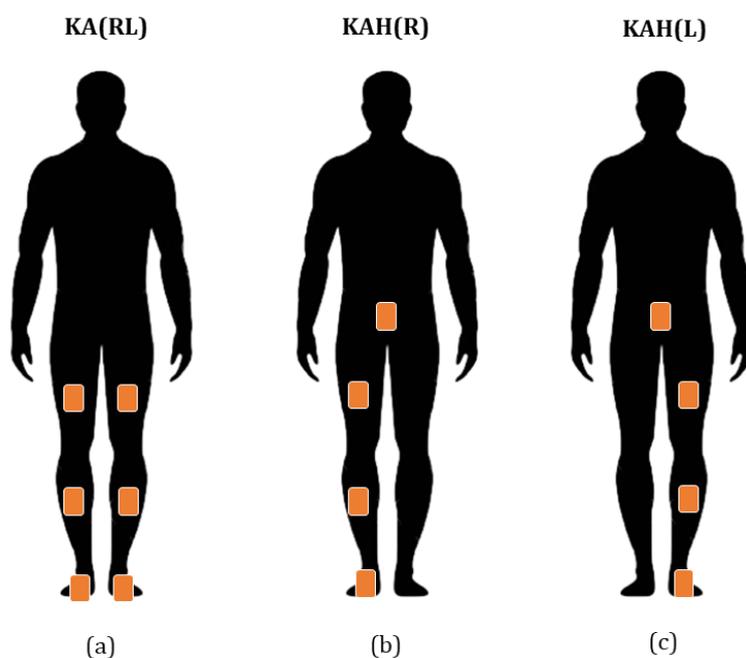


Figura 8.3. Sensores IMU necesarios para recolectar los conjuntos de variables (a) KA(RL), (b) KAH(R) y (c) KAH(L).

Este número es superior al utilizado en Rifaat et al. (2022) o Borzi et al. (2023), estudios en los cuales únicamente se empleó un solo sensor inercial, ubicado en el cinturón de los sujetos. No obstante, tal y como se comentó antes, estos utilizaron como variable la información proporcionada por los ejes del acelerómetro y giroscopio. Esos mismos tipos de entrada fueron los usados en la red CNN-GRU en Martindale et al. (2021), entrenada con los datos recogidos a partir de 4 dispositivos IMU, es decir, el mismo número que este TFG ha concluido como suficiente. Sin embargo, estos solo consideraron dos posibles fases para los pies, postura o swing, no distinguiendo entre las distintas subfases de las mismas, como se ha hecho en el presente trabajo.

El trabajo de Kreuzer & Munz (2021) podría considerarse como el más similar a este TFG. En él, obtuvieron los datos de entrenamiento a partir de once dispositivos IMU e hicieron uso de redes CNN y LSTM para conseguir llevar a cabo la tarea de clasificación. No obstante, consideraron un total de 5 fases de la marcha (contacto inicial, respuesta a la carga, fase media de apoyo, fase final de apoyo y swing), mientras que, para el presente estudio, se decidió limitarse a identificar 4 fases, dado que no sólo se trabajaba con sujeto sano, sino también con paciente con parálisis cerebral. Con todo lo dicho, Kreuzer & Munz (2021) consiguieron un resultado de *accuracy* final de 0.92, es decir, inferior al nuestro, ubicado en torno a los valores 0.97-1 y, todo ello, empleando un menor número de unidades IMU y una arquitectura de red mucho más sencilla.

En segundo lugar, enfocándonos en el Bloque 2, es decir, en la segmentación de la marcha de sujeto sano pediátrico a partir de una red entrenada con sujeto sano adulto, podríamos extraer las siguientes ideas. Por una parte, en vista de los resultados obtenidos en el grupo de test, cabe mencionar que parece que existen diferencias entre la marcha de adultos y la de niños. Es cierto que, analizando con detenimiento las capturas de vídeo de ambos conjuntos poblacionales, se ha podido ver como los movimientos de los sujetos pediátricos, especialmente los de los brazos, se vieron mucho más influenciados por el ambiente de laboratorio, que los de los adultos, cuya marcha podría considerarse prácticamente idéntica a la que hacen cuando no están siendo grabados. Así pues, los conjuntos de *train* y test de este bloque de entrenamiento discrepaban, en pequeños aspectos, entre sí. En consecuencia, en este caso, la sencilla red DNN fue incapaz de captar las particularidades y patrones de los datos, obteniendo unas métricas de rendimiento inferiores a las alcanzadas en el Bloque 1, para los mismos conjuntos de variables. Sin embargo, con el cambio de arquitectura a la red CNN, rápidamente volvieron a conseguirse unos buenos resultados. En cuanto al conjunto óptimo de variables, de nuevo se ha podido confirmar como KA(RL), KAH(R) o KAH(L) son datos más que suficientes para alcanzar un buen rendimiento para la tarea de clasificación. De esta manera, teniendo en cuenta las mismas consideraciones presentadas durante la discusión de los resultados del Bloque 1, este resultado obtenido con datos de vídeo da motivos para pensar que con la arquitectura CNN y los ángulos de flexión tomados a partir de 4-6 sensores IMU es posible hacer una segmentación de la marcha en niños.

Cabe destacar como, en esta ocasión, es notable como las métricas de sensibilidad y F1-Score, tienen un valor, varias centésimas inferior, al obtenido en especificidad y *accuracy* en alguno de los entrenamientos. La causa de ello reside en que, tal y como puede verse en las matrices de confusión y métricas de rendimiento presentadas debajo, la red se confundió más a la hora de predecir la fase '*Final Stance*', lo cual baja la media del rendimiento global. Más concretamente, puede verse que los errores son debidos a que, cuando esa era la clase real, la CNN asignaba a esa muestra un valor de etiqueta correspondiente a la clase '*Medium Stance*'. En consecuencia, el número de falsos negativos fue elevado para esa fase de la marcha, de ahí que disminuyan los valores de sensibilidad y F1-Score, en cálculo aparece el número de falsos negativos en el denominador. En cualquier caso, este comportamiento es algo normal, dado que, si nos fijamos en el número de muestras, esta es la clase con menor representación (es la que dura menos tiempo del ciclo de la marcha), de forma que la red no aprende tan eficientemente sus particularidades como lo hace con el resto de las clases.

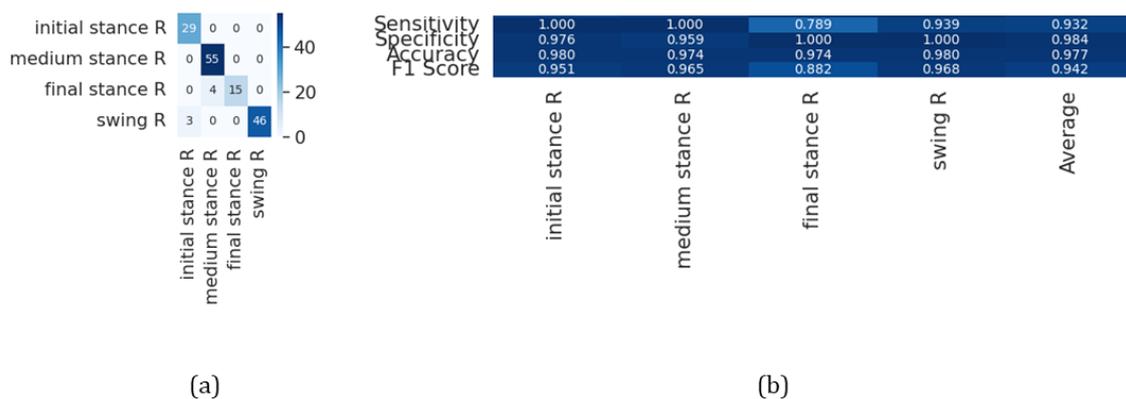


Figura 8.4. Resultados del entrenamiento del clasificador derecho del Bloque 2 utilizando como conjunto de variables KAH(R). (a) Matriz de confusión. (b) Métricas de rendimiento.

Por último, a lo largo de estos dos primeros bloques, se ha podido notar como con los datos de flexión correspondientes al lado izquierdo del cuerpo se han obtenido unos mejores resultados, tanto para el clasificador derecho, como para el izquierdo. Este hecho no se debe a la cantidad existente de ventanas y, por lo tanto, de información disponible, dado que el número de estas es mayor en el lado derecho que en el izquierdo, concretamente 690 frente a 651. De este modo, la razón de esto seguramente se deba a la manera en la que se les indicó a los sujetos que realizasen las capturas de marcha. Siempre se les pidió que empezasen con el pie derecho, por lo que, al tener que arrancar el movimiento con esta pierna se pierde, durante este momento inicial, la fluidez propia del movimiento de marcha. Así pues, si sumamos que cada persona empieza el movimiento de forma diferente y que, debido al reducido espacio, no fue posible dar muchos pasos con cada pie de forma que se eliminase el sesgo de ese primer paso, tiene sentido pensar que los movimientos del lado izquierdo de los distintos sujetos sean más similares entre sí, y, por lo tanto, logren un mejor rendimiento de la red.

8.2 Análisis de resultados en paciente con parálisis cerebral.

Durante el tercer bloque de entrenamientos, se pudo comprobar como para el clasificador derecho, a pesar de entrenar y validar la red con sujetos sanos, las métricas de rendimiento sobre el conjunto de test, compuesto por los sujetos pediátricos con PC, fueron bastante elevadas, prácticamente todas superiores a 0.90. Estos resultados son lógicos, teniendo en cuenta que la mayor parte de los sujetos enfermos, concretamente cinco, tenían afectación en el lado izquierdo y, que los tres restantes, a pesar de tener afectación en el lado derecho, esta no era tan marcada como la vista para el lado izquierdo, donde se podía distinguir claramente el fenómeno de marcha en punta en los correspondientes pacientes. Asimismo, esto explica el que, rompiendo con lo visto anteriormente, con los datos procedentes de movimientos del miembro izquierdo se logre un peor rendimiento del clasificador derecho.

Sin embargo, en el caso del clasificador izquierdo, los resultados fueron muy inferiores a los que se llevaban obteniendo hasta entonces. Ni siquiera eliminando la variable de flexión del tobillo, que se correspondía con la que más difería con respecto al movimiento en sujetos sanos, dada la marcha en punta, consiguió mejorar las métricas. Así pues, en vista de estos datos, es posible afirmar que es necesario crear un sistema específico para clasificar y segmentar la marcha en niños con PC, entrenado únicamente con datos pertenecientes a sujetos con este tipo de condición y no únicamente con sujetos sanos, ya sean niños o adultos.

En cuanto al conjunto de variables disponibles, ha podido verse como la información del tren superior del cuerpo dificulta el aprendizaje de la red, consecuencia de la inexistencia de un braceo fluido en los niños con parálisis cerebral. Además, igual que anteriormente información sobre la flexión de cadera ayuda a mejorar resultados cuando solo se utilizaban los conjuntos KA(R) o KA(L), en esta ocasión ocurría lo contrario.

También, durante el último de los bloques de entrenamientos, pudieron extraerse importantes ideas. Centrándonos en el clasificador derecho y sus resultados, con la arquitectura de red CNN y los conjuntos de variables KA(RL), KAH(R) o KAH(L) puede alcanzarse un rendimiento en validación cruzada prácticamente igual de bueno al mostrado en los Bloques 1 y 2. Es más, ahora de nuevo añadir información acerca de la flexión de cadera consigue mejorar los resultados, a diferencia de lo que ocurría en el Bloque 3. La razón de ello recae en que el movimiento de esta articulación difiere gran cantidad entre sujetos sanos y con PC, sin embargo, dentro de estos segundos, es bastante similar en todos. Siguiendo con este clasificador, si comparamos las matrices de confusión en función de si los datos provienen del lado izquierdo o del derecho, vemos como la clase en la que el porcentaje de error es mayor es 'Initial Stance R' y 'Swing R', en la primera de las situaciones presentadas, y 'Final Stance R' en la restante. Esto tiene sentido, ya que, por una parte, es lógico pensar que con únicamente los ángulos de flexión del lado izquierdo del cuerpo no logremos clasificar correctamente las fases de contacto inicial y swing en el pie derecho, dado que, en los pacientes con afectación en el lado izquierdo del cuerpo, justo eran en estas dos fases donde se daban los eventos más visibles de marcha en punta. Por otro lado, cuando se usa solo como variables información del lado derecho del cuerpo, es normal que la fase peor clasificada sea la de apoyo final, ya que es la que tienen en menor número de ventanas, esto es, una menor representación.

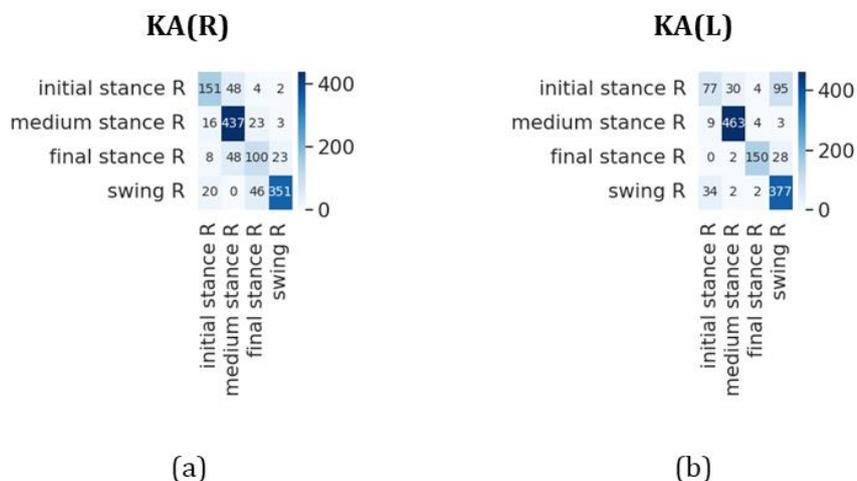


Figura 8.5. Matrices de confusión del clasificador derecho del Bloque 4 utilizando como conjunto de variables: (a) KA(R) y (b) KA(L).

Finalmente, para entrenar al clasificador izquierdo durante el Bloque 4, ha podido verse como ha sido necesario añadir un par de etiquetas nuevas, dado que la afectación en este lado del cuerpo era demasiado notable como para poder considerar como iguales el apoyo inicial o swing en los sujetos con afectación en el lado derecho (los cuales se encontraban más o menos dentro de la normalidad), y esas mismas fases en pacientes con afectación en la mitad izquierda del cuerpo. A pesar de que con esta acción el rendimiento mejoró en gran medida, con ninguno de los conjuntos de variables usados fue posible lograr que las

métricas de sensibilidad y F1-Score superasen el valor 0.5. Estos últimos datos contrastan con los resultados superiores al valor 0.85 para ambas métricas, obtenidos en el estudio de Borzì et al. (2023), el cual también trabajó con sujetos con afectación en la marcha, en este caso, con Enfermedad del Parkinson (EP). Es cierto que, en este último estudio, la tarea únicamente era detectar eventos de congelación de la marcha, es decir, no hacía una segmentación de la marcha en 4 fases, como ha sido el objetivo de este TFG, lo que hace mucha más sencillo el trabajo de clasificación de Borzì et al. (2023). Sin embargo, podría afirmarse que la principal causa del bajo rendimiento del clasificador izquierdo del Bloque 4 es el escaso *dataset* disponible. La CNN solamente ha podido extraer patrones y características de un total de 8 pacientes con PC, mientras que este número asciende al de 128 sujetos con EP en el otro estudio (Borzì et al., 2023). La baja disponibilidad de datos de entrenamiento puede comprobarse en la Figura 8.6, donde se observa como las fases ‘*PC swing L*’ y ‘*PC initial stance L*’ prácticamente son siempre clasificadas erróneamente. La mayoría de estos fallos consisten en que estas son confundidas por sus respectivas fases no patológicas, esto es, ‘*Swing L*’ e ‘*Initial Stance L*’. Al fin y al cabo, estos resultados no son de extrañar. De por sí 8 sujetos ya es un número reducido de *dataset*, pero, además, en este caso, solamente son 4 los que contienen las etiquetas 5-6 en sus muestras y, de ellos, únicamente 2 han sido empleados para entrenar la red.

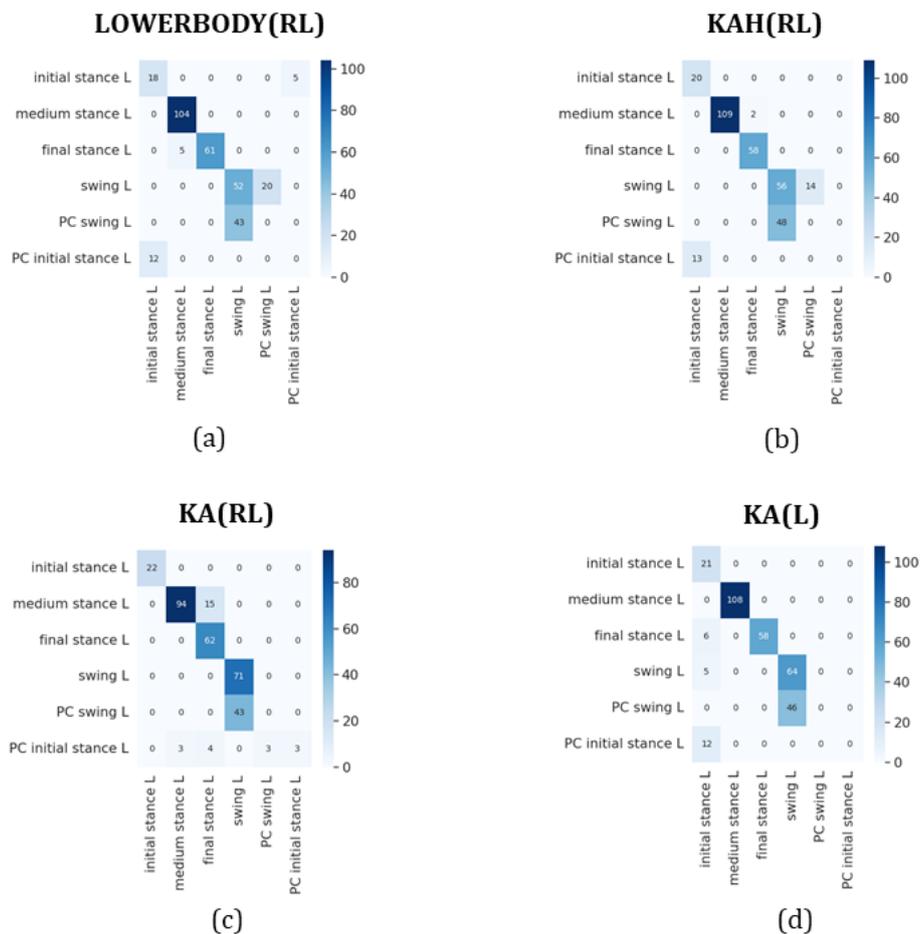


Figura 8.6. Matrices de confusión del clasificador izquierdo del Bloque 4 utilizando etiquetas 1-6 y como conjunto de variables: (a) LOWERBODY(RL), (b) KAH(RL), (c) KA(RL) y (d) KA(L).

8.3 Limitaciones

Mientras se desarrolló este estudio, se encontraron algunas limitaciones y/o aspectos mejorables. En primer lugar, tal y como se ha comentado con anterioridad, el número de sujetos empleados como *dataset* es relativamente bajo, contando con tan solo 14 sujetos adultos sanos, 3 sujetos pediátricos sanos y 8 sujetos pediátricos con parálisis cerebral. Esto limita y dificulta la capacidad de aprendizaje de la red, ya que, con una muestra tan escasa, es imposible que las distintas arquitecturas de deep learning creadas sean capaces de extraer patrones de los datos, así como generalizar. Una solución sería aumentar el número de instancias e igualar el número de sujetos en los tres grupos. De este modo, sería posible incrementar el número de entrenamientos de validación cruzada y se conseguiría que la base de datos estuviese más balanceada.

Por otra parte, el espacio donde se realizaron las capturas de datos era muy limitado, lo que imposibilitó hacer más largas las grabaciones en que se utilizó *OpenCap* y *Xsens MVN Awiinda* de forma conjunta, debido a que, el primero de estos dos métodos requiere que el sujeto siempre este siendo visionado por los dos dispositivos *smartphones*. Como consecuencia, esto también afecta al reducido tamaño de la base de datos recopilada, ya que, únicamente fue posible registrar 4-5 pasos para cada sentido de marcha en cada uno de los sujetos. Asimismo, las reducidas dimensiones de la Sala Leonardo Torres Quevedo impidieron la captura de marcha larga con los dispositivos Xsens, lo que implica que en sujetos adultos de control no se pudo obtener el análisis de la marcha en MotionCloud. Así pues, en futuros estudios, sería conveniente hacer las capturas en espacios mucho más amplios.

Para finalizar este apartado, sería necesario comentar la tarea de etiquetado. Esta se ha hecho de forma visual, sin ayuda de ningún profesional del ámbito sanitario especializado en la marcha. De esta forma, es posible que exista cierta subjetividad que haya podido influir en los resultados obtenidos en este trabajo de fin de grado. Hubiese sido necesario hacer una validación y comparación con datos etiquetados por expertos sanitarios, sin embargo, por cuestiones de tiempo, esta prueba no ha podido ser realizada.

9. CONCLUSIONES Y LÍNEAS FUTURAS

9.1 Cumplimiento de los objetivos del trabajo de fin de grado

Para la elaboración de este TFG, se han desarrollado un conjunto de redes de aprendizaje profundo para realizar una segmentación de la marcha en sus respectivas fases, tanto en sujetos de control como con parálisis cerebral. De este modo, se ha conseguido mostrar el potencial de dispositivos económicos, con los que es posible elaborar y generalizar el uso del análisis instrumentado de la marcha en los hospitales. En el Capítulo 1. Introducción, se enumeraron un conjunto de objetivos específicos a cumplir durante la realización de este trabajo. En consecuencia, se va a valorar el cumplimiento de estos:

- i. Se han revisado varios artículos, libros y tesis relacionados con el análisis instrumentado de la marcha en la parálisis cerebral, así como otros que abordaban tareas de segmentación de la marcha. A través de esta lectura, se ha conocido el estado actual en el que se haya el análisis instrumentado de la marcha y se han descubierto los tipos de redes más empleados para abordar la clasificación de la marcha, sirviendo como inspiración al presente trabajo.
- ii. Se ha capturado una base de datos multimodales de marcha en sujeto sano y paciente con parálisis cerebral, empleando simultáneamente el sistema de vídeo OpenCap y las unidades de medición inercial Xsens.
- iii. Se ha desarrollado una interfaz en Python con la que se ha logrado agilizar la tarea de etiquetado manual de las fases de la marcha, empleando como datos los vídeos capturados mediante OpenCap.
- iv. Se han creado dos arquitecturas de red en deep learning, con las que se han ejecutado varios bloques de entrenamientos, utilizando en cada uno de ellos una combinación distinta de los 3 grupos disponibles en el *dataset*.
- v. Se han evaluado los resultados obtenidos y se han contrastado con estudios previos para verificar la coherencia de dichos resultados.
- vi. Tras discutir los hallazgos de este trabajo, se exponen un conjunto de conclusiones alcanzadas junto con las limitaciones encontradas, así como posibles direcciones para futuras investigaciones.

9.2 Conclusiones

Una vez discutidos los resultados obtenidos, a continuación, se van a enumerar las conclusiones extraídas. Estas son las siguientes:

1. Para la clasificación de la marcha en sujeto sano, dados los resultados obtenidos a partir de imágenes de vídeo, resulta factible pensar que es posible alcanzar una óptima tarea de segmentación empleando únicamente 4 o 6 sensores inerciales. De esta manera, se reduciría el tiempo requerido para la colocación de los mismos. Asimismo, los resultados alcanzados apoyan la idea de que empleando como conjunto de variables la flexión de rodilla y tobillo en ambos miembros inferiores, o bien, la flexión de cadera, rodilla y tobillo de uno solo de los miembros, es suficiente para lograr un buen aprendizaje de la red.
2. Estructuras simples de red, como la DNN o CNN desarrolladas, son capaces de extraer patrones de los datos y lograr un buen rendimiento, consiguiendo así

entrenamientos más rápidos y computacionalmente menos costosos que arquitecturas de mayor profundidad.

3. Al segmentar la marcha en paciente con PC, habiendo entrenado previamente la red con paciente sano, se ha visto como es más adecuado no considerar las variables del tren superior del cuerpo. Asimismo, aparentemente el movimiento de cadera también difiere notablemente entre ambos tipos de sujetos.
4. Al clasificar a sujetos pediátricos con parálisis cerebral, se observa cómo es necesario desarrollar un modelo de clasificación entrenado de forma independiente con pacientes con esta condición, necesitando de la creación de nuevas etiquetas cada vez que se dé de algún tipo de desviación respecto a los movimientos normales de cada una de las fases de la marcha.
5. Con ambas herramientas presentadas, es posible realizar un análisis instrumentado de la marcha que sirva como apoyo al análisis visual realizado por el médico. De esta manera, se obtiene un conjunto de datos más preciso y objetivo, así como un tratamiento más individualizado para cada paciente. Adicionalmente, el reducido precio de ambos sistemas, en comparación con el actual gold standard Vicon, hace posible la extensión de este tipo de análisis en más centros hospitalarios.

9.3 Líneas futuras

Como se mencionó previamente, una línea prometedora para el futuro sería ampliar el tamaño de la base de datos, al incluir un mayor número de sujetos, especialmente del grupo pediátrico, tanto sanos como enfermos, para conseguir equilibrar las distintas clases de sujetos.

En segundo lugar, para el desarrollo de la red únicamente se ha empleado la base de datos correspondiente a OpenCap. Así pues, sería interesante comprobar si empleando como entrada las variables recogidas a partir de sensores IMU, el rendimiento de esas mismas arquitecturas neuronales se mantiene, mejora o empeora. Una manera de verificar esto, sería desarrollando un código en Python, el cual, a partir de las etiquetas asignadas a cada muestra, fuese capaz de graficar la evolución de los ángulos de las articulaciones a lo largo del ciclo. De esta manera, se podría examinar si existe similitud entre estas gráficas, obtenidas a partir del etiquetado de los datos recogidos por Xsens, y las gráficas que aporta el procesado de estos mismos datos en MotionCloud.

Por último, ha podido verse como en ciertos artículos científicos se empleaban redes mucho más profundas y complejas para la segmentación de las fases de la marcha. Resultaría interesante comprobar si con este tipo de arquitecturas y la actual base de datos adquirida en este trabajo, es posible alcanzar mejores resultados.

BIBLIOGRAFÍA

- Armand, S., Decoulon, G., & Bonnefoy-Mazure, A. (2016). Gait analysis in children with cerebral palsy. *EFORT Open Reviews*, *1*(12), 448-460. <https://doi.org/10.1302/2058-5241.1.000052>
- Bekteshi, S., Monbaliu, E., McIntyre, S., Saloojee, G., Hilberink, S. R., Tatishvili, N., & Dan, B. (2023). Towards functional improvement of motor disorders associated with cerebral palsy. *The Lancet Neurology*, *22*(3), 229-243. [https://doi.org/10.1016/S1474-4422\(23\)00004-2](https://doi.org/10.1016/S1474-4422(23)00004-2)
- Blair, S., Duthie, G., Robertson, S., Hopkins, W., & Ball, K. (2018). Concurrent validation of an inertial measurement system to quantify kicking biomechanics in four football codes. *Journal of Biomechanics*, *73*, 24-32. <https://doi.org/10.1016/j.jbiomech.2018.03.031>
- Borzi, L., Sigcha, L., Rodríguez-Martín, D., & Olmo, G. (2023). Real-time detection of freezing of gait in Parkinson's disease using multi-head convolutional neural networks and a single inertial sensor. *Artificial Intelligence in Medicine*, *135*, 102459. <https://doi.org/10.1016/j.artmed.2022.102459>.
- Bosch Rué, A. (2019). *Deep learning: Principios y fundamentos / Anna Bosch Rué, Jordi Casas-Roma, Toni Lozano Bagén*. (Primera edición en lengua castellana). Editorial UOC.
- Gait Report—How to interpret data?* (s. f.). Recuperado 16 de mayo de 2023, de https://base.xsens.com/s/article/Gait-Report-How-to-interpret-data?language=en_US
- Gao, S., Chen, J., Dai, Y., & Hu, B. (2022). *Wearable Systems Based Gait Monitoring and Analysis*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-97332-2>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning* (Illustrated edition). The MIT Press.
- Haakana, P., Nurminen, J., Kulmala, J. P., Niemelä, T., Marttinen Rossi, E., Mäenpää, H., & Piitulainen, H. (2020). Effects of cognitive and motor dual-task on spatiotemporal gait parameters in children and adolescents with Cerebral Palsy. *Gait & Posture*, *81*, 130-131. <https://doi.org/10.1016/j.gaitpost.2020.07.099>
- Hebda-Boon, A., Tan, X.-L., Tillmann, R., Shortland, A. P., Firth, G. B., & Morrissey, D. (2023). The impact of instrumented gait analysis on decision-making in the interprofessional management of cerebral palsy: A scoping review. *European Journal of Paediatric Neurology*, *42*, 60-70. <https://doi.org/10.1016/j.ejpn.2022.11.007>
- Jankovic, J., PhD, J. C. M. M., PhD, S. P. M., & MD, N. J. N. (2022). *Bradley y Daroff. Neurología clínica*.
- Keras: Deep Learning for humans*. (s. f.). Recuperado 30 de mayo de 2023, de <https://keras.io/>
- Kharb, A., Saini, V., Jain, Y., Dhiman, S., Tech, M., & Scholar. (2011). A review of gait cycle and its parameters. *IJCEM Int J Comput Eng Manag*, *13*.
- Kidziński, Ł., Delp, S., & Schwartz, M. (2019). Automatic real-time gait event detection in children using deep neural networks | *PLOS ONE*. <https://doi.org/10.1371/journal.pone.0211466>

Kreuzer, D., & Munz, M. (2021). Deep Convolutional and LSTM Networks on Multi-Channel Time Series Data for Gait Phase Recognition. *Sensors*, 21(3), Article 3. <https://doi.org/10.3390/s21030789>

Kumar, V. S., Banjara, R., Kapoor, L., Gupta, M., Aggarwal, P., Rano, D. D., & Khan, S. A. (2021). Does instrumented gait analysis help decision making in Cerebral palsy treatment? - Early experience from a developing country. *Gait & Posture*, 90, 223. <https://doi.org/10.1016/j.gaitpost.2021.09.115>

La Función de Activación. (s. f.). Codificando Bits. Recuperado 30 de abril de 2023, de <https://www.codificandobits.com/blog/funcion-de-activacion/>

López-Nava, I., & Muñoz-Meléndez, A. (2016). Wearable Inertial Sensors for Human Motion Analysis: A Review. *IEEE Sensors Journal*, PP, 1-1. <https://doi.org/10.1109/JSEN.2016.2609392>

Martindale, C. F., Christlein, V., Klumpp, P., & Eskofier, B. M. (2021). Wearables-based multi-task gait and activity segmentation using recurrent neural networks. *Neurocomputing*, 432, 250-261. <https://doi.org/10.1016/j.neucom.2020.08.079>

Mitchell, T. (1997). *Machine Learning*.

Nowicki, P. D. (Ed.). (2020). *Orthopedic Care of Patients with Cerebral Palsy: A Clinical Guide to Evaluation and Management across the Lifespan*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-46574-2>

Papageorgiou, E., Nieuwenhuys, A., Vandekerckhove, I., Van Campenhout, A., Ortibus, E., & Desloovere, K. (2019). Systematic review on gait classifications in children with cerebral palsy: An update. *Gait & Posture*, 69, 209-223. <https://doi.org/10.1016/j.gaitpost.2019.01.038>

Parmar, R. (2018, septiembre 11). *Training Deep Neural Networks*. Medium. <https://towardsdatascience.com/training-deep-neural-networks-9fdb1964b964>

Perry, J., K, S. T., & Davids, J. R. (1992). Gait Analysis: Normal and Pathological Function. *Journal of Pediatric Orthopaedics*, 12(6), 815.

Prieto, Á. G. (2016). *Niños y niñas con parálisis cerebral: Descripción, acción educativa e inserción social*. Narcea Ediciones. <https://elibro-net.ponton.uva.es/es/ereader/uva/45973>

Richards, C. L., & Malouin, F. (2013). Chapter 18 - Cerebral palsy: Definition, assessment and rehabilitation. En O. Dulac, M. Lassonde, & H. B. Sarnat (Eds.), *Handbook of Clinical Neurology* (Vol. 111, pp. 183-195). Elsevier. <https://doi.org/10.1016/B978-0-444-52891-9.00018-X>

Rifaat, N., Ghosh, U. K., & Sayeed, A. (2022). Accurate gait recognition with inertial sensors using a new FCN-BiLSTM architecture. *Computers and Electrical Engineering*, 104, 108428. <https://doi.org/10.1016/j.compeleceng.2022.108428>

Song, C., Huang, Y., Huang, Y., Jia, N., & Wang, L. (2019). GaitNet: An end-to-end network for gait based human identification. *Pattern Recognition*, 96, 106988. <https://doi.org/10.1016/j.patcog.2019.106988>

States, R. A., Krzak, J. J., Salem, Y., Godwin, E. M., Bodkin, A. W., & McMulkin, M. L. (2021). Instrumented gait analysis for management of gait disorders in children with cerebral palsy: A scoping review. *Gait & Posture*, *90*, 1-8. <https://doi.org/10.1016/j.gaitpost.2021.07.009>

Tugui, R. D., & Antonescu, D. (2013). Cerebral Palsy Gait, Clinical Importance. *Mædica*, *8*(4), 388-393.

Tunca, C., Pehlivan, N., Ak, N., Arnrich, B., Salur, G., & Ersoy, C. (2017). Inertial Sensor-Based Robust Gait Analysis in Non-Hospital Settings for Neurological Disorders. *Sensors*, *17*(4), Article 4. <https://doi.org/10.3390/s17040825>

Uhlrich, S. D., Falisse, A., Kidziński, Ł., Muccini, J., Ko, M., Chaudhari, A. S., Hicks, J. L., & Delp, S. L. (2022). *OpenCap: 3D human movement dynamics from smartphone videos* (p. 2022.07.07.499061). bioRxiv. <https://doi.org/10.1101/2022.07.07.499061>

Wren, T. A. L., Otsuka, N. Y., Bowen, R. E., Scaduto, A. A., Chan, L. S., Sheng, M., Hara, R., & Kay, R. M. (2011). Influence of gait analysis on decision-making for lower extremity orthopaedic surgery: Baseline data from a randomized controlled trial. *Gait & Posture*, *34*(3), 364-369. <https://doi.org/10.1016/j.gaitpost.2011.06.002>

Xsens—MVN Gait report white paper.pdf. (s. f.). Recuperado 30 de abril de 2023, de <https://www.xsens.com/hubfs/Downloads/Whitepapers/Xsens%20-%20MVN%20Gait%20report%20white%20paper.pdf>