

Universidad San Jorge

Escuela de Arquitectura y Tecnología

Grado en Ingeniería Informática

Proyecto Final

**Desarrollo de un sistema de reconocimiento de
poses 3D en personas utilizando ML**

Autor del proyecto:

Marcos Marina Castelló

Directores académicos:

Violeta Monasterio Bazán y Alfonso Santolaria Lorenzo

Zaragoza, 30 de septiembre de 2023



Este trabajo constituye parte de mi candidatura para la obtención del título de Graduado en Ingeniería Informática por la Universidad San Jorge y no ha sido entregado previamente (o simultáneamente) para la obtención de cualquier otro título.

Este documento es el resultado de mi propio trabajo, excepto donde de otra manera esté indicado y referido.

Doy mi consentimiento para que se archive este trabajo en la biblioteca universitaria de Universidad San Jorge, donde se puede facilitar su consulta.

Firma

A handwritten signature in black ink, appearing to read 'Walter', with a stylized flourish above the name.

Fecha

12 septiembre 2023

Dedicatoria y Agradecimiento

El desarrollo de este proyecto ha sido el resultado de un esfuerzo conjunto en el que diversas personas han tenido una participación activa, ya sea de manera directa o indirecta, ofreciendo sus opiniones, sugerencias, correcciones, y brindando su respaldo y paciencia. Gracias a esta colaboración he podido aprovechar el conocimiento y experiencia de muchas personas, a quienes deseo manifestar mi profundo agradecimiento en este apartado.

A ITAINNOVA, por acceder a realizar el primer contrato en alternancia con la universidad con el objetivo de desarrollarme profesionalmente, así como a la propia Universidad San Jorge por su predisposición para mi incorporación en esta modalidad.

A mis tutores académicos, Violeta Monasterio Bazán y Alfonso Santolaria Lorenzo, por su orientación y apoyo durante el proceso de elaboración del PFG. Sin su ayuda, no habría sido posible alcanzar los objetivos propuestos.

A mis tutores profesionales, Vega Rodríguez Chamarro y Rafael del Hoyo Alonso, por su asesoramiento, consejo y respaldo. Quiero hacer especial mención a Rafael por su gran entrega y diligencia en el desarrollo de este proyecto. Su valiosa experiencia y habilidades de gestión de proyectos han sido fundamentales para llevar a cabo este trabajo con éxito. Estoy sinceramente agradecido por su compromiso y apoyo constante.

A mi familia, en especial a mis padres quienes han sido mi mayor fuente de inspiración, apoyo y amor incondicional a lo largo de mi vida. Ellos me han enseñado el valor del trabajo duro, la perseverancia, el respeto y la honestidad, y gracias a su dedicación y sacrificio he podido llegar hasta aquí.

Por último, pero no menos importante, quiero agradecer a todas las personas que, aunque no hayan sido mencionadas anteriormente, han formado parte de mi camino y me han brindado su ayuda de una forma u otra. Desde mi novia a mis compañeros de departamento (en especial Ana Hernández Ruiz, Carlos Marañes Nueno e Irene López Bosque que me han guiado en mi día a día), así como también a mis profesores, compañeros de clase, amigos, compañeros de equipo, conocidos y extraños, pues cada una de estas personas ha contribuido de alguna forma en mi formación y crecimiento personal.

Espero que esta dedicatoria sea un reflejo de mi gratitud y reconocimiento a cada una de las personas que han estado a mi lado y me han ayudado a alcanzar este logro.

Tabla de contenido

Resumen	1
Abstract	1
Palabras clave	2
1. Introducción.....	3
2. Estado del arte y antecedentes.....	5
2.1. Estado del arte	5
2.1.1. Estimación de Pose Humana (HPE- Human Pose Estimation)	5
2.1.2. Reconocimiento de actividades humanas con IA (HAR-Human Activity Recognition)	12
2.2. Antecedentes.....	14
3. Objetivos	19
4. Metodología	21
4.1. Bases del proyecto	21
4.2. ¿Por qué Scrumban?.....	22
4.2.1. Scrum en Scrumban.....	23
4.2.2. Kanban en Scrumban.....	24
4.3. Gestión de versiones	26
4.3.1. GitLab y GitHub.....	26
4.4. Estrategias de comunicación.....	26
5. Estudio, diseño e implementación	29
5.1. Sprint 1: Estimación de Pose, ExPose, Bodyflow	29
5.1.1. Configuración del tablero Kanban	29
5.1.2. Tarea 1- Aprendizaje en conceptos básicos y funcionamiento red neuronal.....	29
5.1.3. Tarea 2- Estudio SMPL.....	30
5.1.4. Tarea 3- Estudio SMPL-X.....	31
5.1.5. Tarea 4- Estudio ExPose.....	31
5.1.6. Tarea 5- Comprensión utilidad y usabilidad Bodyflow	33
5.1.7. Evaluación del sprint.....	33
5.2. Sprint 2: Preparación del entorno y análisis del estado actual.....	34
5.2.1. Configuración tablero Kanban	34
5.2.2. Tarea 1- Instalación y configuración Visual Studio Code – servidores ITAINNOVA	34
5.2.3. Tarea 2- Creación de ambientes de Conda.....	35
5.2.4. Tarea 3- Elaboración del Diagrama de clases para Bodyflow	35
5.2.5. Tarea 4- Elaboración del Diagrama de clases para ExPose.....	37

5.2.6.	Tarea 5- Analizar ExPose para comprender su funcionamiento y exponerlo	37
5.2.7.	Tarea 6- Realización de una Demo de ExPose	39
5.2.8.	Evaluación del sprint.....	40
5.3.	<i>Sprint 3: Integración de ExPose y visualización de resultados</i>	41
5.3.1.	Configuración tablero Kanban	41
5.3.2.	Tarea 0- Analizar ExPose para comprender su funcionamiento y exponerlo	41
5.3.3.	Tarea 1 - Estudio de las licencias de ExPose	44
5.3.4.	Tarea 2- Redacción y exposición de la planificación de integración.....	45
5.3.5.	Tarea 3- Implementación	46
5.3.6.	Tarea 4- Creación de un wrapper (envoltorio) que facilite la instalación	46
5.3.7.	Tarea 5- Corrección de paths propio de este tipo de integración	46
5.3.8.	Tarea 6- Comprobación de funcionamiento a través de tests unitarios	47
5.3.9.	Tarea 7- Profiling (perfilado) del código	47
5.3.10.	Tarea 8- Gráfico del flujo de funcionamiento de ExPose tras su integración	48
5.3.11.	Tarea 9- Diseño de los formatos de salida de los datos generados.....	49
5.3.12.	Tarea 10- Visualización de la salida a través de un vídeo integrado	51
5.3.13.	Evaluación del sprint.....	51
5.4.	<i>Sprint 4: Módulo de sincronización</i>	52
5.4.1.	Configuración tablero Kanban	52
5.4.2.	Tarea 1- Análisis del funcionamiento módulo de sincronización	52
5.4.3.	Tarea 2- Adaptar el funcionamiento a la nueva funcionalidad multipose	53
5.4.4.	Tarea 3- Adaptar al cambio en la denominación de los tiempos (timestamps)	53
5.4.5.	Tarea 4- Adaptar ExPose para su funcionamiento en serie.....	53
5.4.6.	Tarea 5- Ejecución y sincronización del "FALL-UP dataset" con ExPose.....	53
5.4.7.	Evaluación del sprint.....	54
5.5.	<i>Sprint 5: Evaluación del modelo y resultados</i>	55
5.5.1.	Configuración tablero Kanban	55
5.5.2.	Tarea 1- Análisis del funcionamiento módulo de HAR	55
5.5.3.	Tarea 2- Ejecutar las diferentes combinaciones de modelos HAR y características que sean útiles para la evaluación de ExPose.....	55
5.5.4.	Tarea 3- Presentación y análisis de resultados.....	57
5.5.5.	Evaluación del sprint.....	57
6.	<i>Estudio económico</i>	59
6.1.	<i>Costes económicos</i>	59
6.1.1.	Costes materiales	59
6.1.2.	Costes humanos	60

6.1.3.	Costes de desarrollo	61
6.1.4.	Costes Totales.....	62
6.2.	Beneficios económicos	62
7.	<i>Resultados</i>	63
7.1.	Desarrollo del proyecto con Scumban	63
7.2.	Análisis de los resultados obtenidos por la integración de ExPose.....	66
8.	<i>Conclusiones</i>	71
8.1.	Extensiones y mejoras futuras	72
9.	<i>Bibliografía</i>	73
<i>ANEXO I: Propuesta de Proyecto</i>		77
<i>ANEXO II: Actas de reunión</i>		79
Reunión 1		79
Reunión 2		80
Reunión 3		81
Reunión 4		82
Reunión 5		83
Reunión 6		84
Reunión 7		85
Reunión 8		86
Reunión 9		87
Reunión 10.....		88
Reunión 11.....		89
Reunión 12.....		90
Reunión 13.....		91
<i>ANEXO III: Resultados ITAINNOVA</i>		93
<i>ANEXO IV: Material Complementario</i>		95

Resumen

El objetivo principal de este proyecto es lograr la integración de un módulo de estimación de pose 3D dentro de una librería de análisis del movimiento desarrollada por el Instituto Tecnológico de Aragón (ITAINNOVA). El módulo realizará la estimación de pose de manera robusta y rápida utilizando técnicas de Aprendizaje Máquina. Para obtener una comprensión completa de la apariencia, interacción y desempeño de las personas, es crucial capturar de forma rápida y precisa la información tridimensional de su cuerpo, rostro y manos a partir de una imagen en color RGB. La identificación de pose surge de la necesidad de conseguir que los usuarios de sistemas de realidad virtual y aumentada puedan interactuar con estos de una forma más natural y se puedan identificar eventos o determinadas actividades en vídeos. Se trata de un proyecto de investigación y desarrollo con aplicaciones en el ámbito clínico y deportivo.

A lo largo del documento se presenta de manera detallada el proceso de investigación realizado, así como las nuevas aportaciones al proyecto "Misiones" en ITAINNOVA. El resultado ha sido satisfactorio, se ha conseguido integrar el nuevo modelo volumétrico ExPose en el paquete Human Pose Estimation (HPE) de la librería BodyFlow, así como utilizar las métricas para comparar este modelo con los ya existentes en el Human Activity Recognition (HAR) obteniendo valiosos resultados.

Abstract

The main objective of this project is to achieve the integration of a 3D pose estimation module within a motion analysis library developed by the Aragón Institute of Technology (ITAINNOVA). The module will perform robust and fast pose estimation using Machine Learning techniques. To fully understand the appearance, interaction, and performance of individuals, it is crucial to capture the three-dimensional information of their body, face, and hands quickly and accurately from an RGB color image. The need for pose identification arises from the necessity to enable users of virtual and augmented reality systems to interact with them in a more natural way, and to identify events or specific activities in videos. This is a research and development project with applications in the clinical and sports fields.

Throughout the document, the research process is presented in detail, as well as the new contributions to the "Misiones" project at ITAINNOVA. The outcome has been successful; the new volumetric model, ExPose, has been integrated into the Human Pose Estimation (HPE) package of the BodyFlow library, and metrics have been used to compare this model with the existing ones in Human Activity Recognition (HAR), yielding valuable results.

Palabras clave

Inteligencia Artificial, Detección de Pose, Visión Computacional, Aprendizaje Automático, Representación Tridimensional, Reconocimiento de Actividad, Entrenamiento Supervisado

1. Introducción

La detección de pose se refiere a la tarea de estimar la posición y orientación de las articulaciones del cuerpo humano en una imagen o secuencia de imágenes. Es un área de investigación en visión por computadora y aprendizaje automático que tiene aplicaciones en una amplia variedad de campos, como la realidad virtual, la realidad aumentada, la robótica, la monitorización de la salud y el análisis de movimiento humano.

El objetivo principal de la detección de pose es comprender y capturar la postura y movimiento del cuerpo humano. Esto implica identificar y localizar las articulaciones clave del cuerpo, como las muñecas, los codos, las rodillas y los tobillos, así como estimar las conexiones y relaciones entre ellas. Esta información de pose puede utilizarse para rastrear y analizar movimientos humanos, reconocer acciones y gestos, generar animaciones realistas, controlar dispositivos mediante gestos y muchas otras aplicaciones. La detección de pose es un desafío técnico debido a la variabilidad en la apariencia y forma del cuerpo humano, la oclusión parcial de las articulaciones, la presencia de diferentes posturas y la complejidad de los movimientos. Para abordar este desafío, se han desarrollado diversos enfoques y técnicas, que van desde métodos basados en características y modelos estadísticos hasta enfoques basados en aprendizaje profundo y redes neuronales convolucionales (*ver capítulo 2*)

La mayoría de los métodos existentes se centran únicamente en partes específicas del cuerpo. Estos métodos se basan en técnicas de optimización, lo que los hace lentos, propensos a óptimos locales y requieren puntos clave en 2D como entrada. Para abordar estas limitaciones, Vasileios Choutas et al. [41, 42] desarrollaron ExPose (EXpressive POse and Shape rEgression), que realiza directamente la regresión del cuerpo, el rostro y las manos en formato SMPL-X a partir de una imagen en color RGB. El proyecto tiene como objetivo principal lograr la integración de ExPose en la librería BodyFlow de ITAINNOVA. Esta integración permitirá no solo llevar a cabo un proceso de divulgación y presentación de los resultados, sino también comparar el rendimiento de ExPose con los métodos existentes en la librería y evaluar su utilidad práctica.

Este proyecto se encuentra enmarcado en la línea de investigación del grupo de Inteligencia Artificial, Sistemas Cognitivos y Big Data, y dentro del proyecto AI4HEALTHYAGING «IA para el diagnóstico y tratamiento temprano de enfermedades con gran prevalencia en envejecimiento». Programa Misiones de I+D en Inteligencia Artificial, referencia MIA.2021.M02.0007, cuyo objetivo es la investigación y el desarrollo de un sistema inteligente descentralizado de toma y análisis de datos que permita a nuestro sistema de salud anticipar y actuar de manera temprana y ultra rápida en enfermedades psiquiátricas, neurológicas, motoras y degenerativas de gran prevalencia en el contexto del envejecimiento.

La innovación radica en la implementación y aplicación práctica de ExPose, en lugar de limitarse a la investigación teórica. Se facilitará su uso y permitirá a los usuarios aprovecharse de las ventajas que ofrece este enfoque avanzado de detección de pose beneficiándose de un entorno establecido y robusto que ofrece herramientas y recursos adicionales.

2. Estado del arte y antecedentes

En el primer apartado de este capítulo, se abordarán los fundamentos esenciales y las tecnologías previamente establecidas en relación con la estimación de pose. Se explorarán los conceptos clave y las técnicas ya existentes en este campo. Además, se examinará cómo se ha abordado históricamente la estimación de la postura en diversos contextos.

La segunda parte del capítulo se centrará en la presentación de la librería desarrollada por ITAINNOVA, junto con su estado inicial en el momento en que se inició este PFG. Se ofrecerá una visión detallada de la misma y su funcionamiento inicial. Esto proporcionará un punto de referencia crucial para comprender el avance y las mejoras que se lograrán a lo largo del proyecto.

2.1. Estado del arte

2.1.1. Estimación de Pose Humana (HPE- Human Pose Estimation)

Cuando nos referimos a la estimación de la pose humana, hablamos de la tarea de capturar y analizar las articulaciones y el movimiento del cuerpo utilizando técnicas de visión por computadora. Utilizando una imagen o una secuencia de imágenes, se busca estimar la configuración de las diferentes partes del cuerpo humano. Esta área de investigación se encuentra en constante crecimiento debido a sus numerosas aplicaciones y su complejidad.

Como punto de partida en el campo de la visión por computadora aplicada a la estimación de pose humana, los estudios realizados por Gavrilu (1999) [6] son una referencia importante. Por otro lado, los trabajos de Moeslund et al. (2006) [7] y Poppe (2007) [8] así como el estudio realizado por Sarafianos et al. (2016) [5] abordan ampliamente los métodos de visión por computadora utilizados para capturar el movimiento humano, aunque no se centran específicamente en la estimación en 3D y pueden considerarse algo desactualizados en un ámbito que avanza a la velocidad de la luz.

Desde una perspectiva práctica (*Figura 1*), las soluciones ofrecidas por la visión por computadora son altamente atractivas, ya que proporcionan métodos no invasivos para una amplia gama de aplicaciones. Por ejemplo, en el ámbito de la comunicación entre humanos y ordenadores o robots, el reconocimiento de gestos como comandos puede facilitar la interacción [9]. En el mundo de los videojuegos, se utiliza ampliamente esta tecnología, tanto en forma de controles para los jugadores [10] como en la generación de avatares en 3D [11]. En los campos deportivo y médico, encontramos herramientas de análisis y estudio del comportamiento anatómico del cuerpo humano [12]. Además, se han realizado investigaciones para comprender el comportamiento y las relaciones humanas, cuya aplicación puede ser útil en tareas de videovigilancia [13].

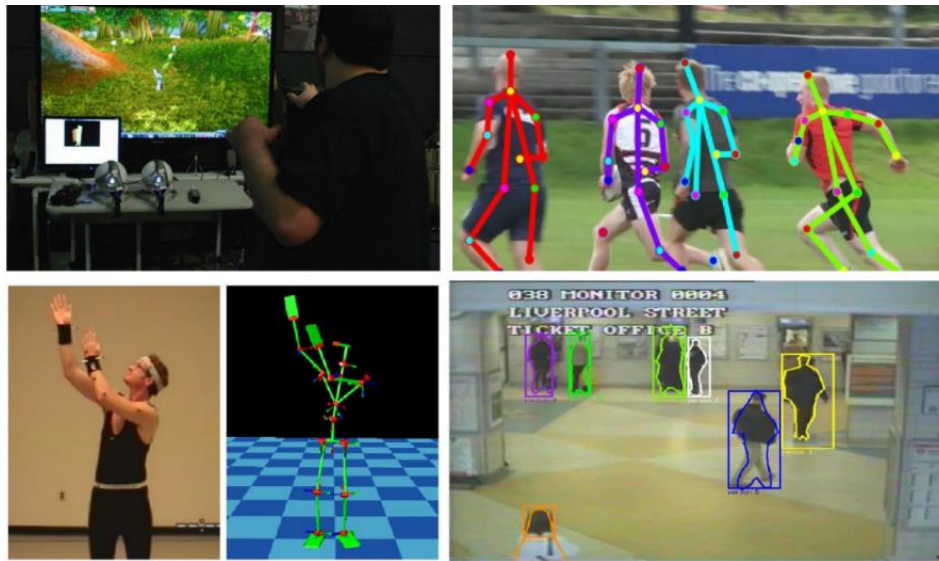


Figura 1: Ejemplos de aplicaciones de la estimación de la pose. De izquierda a derecha y de arriba abajo: control de videojuegos, deportes, creación de animaciones para videojuegos, videovigilancia (Fuente: Suma et al. (2011), Iqbal, Milan y Gall (2017), Zheng y Yamane (2015), Satpathy, Siebel y Rodríguez (2004)) [2]

Debido a que la estimación de pose monocular (MHPE- *Monocular Human Pose Estimation*) no requiere cámaras múltiples o marcadores portátiles, se ha convertido en una parte significativa de muchas aplicaciones del mundo real, como la realidad virtual, la producción y edición de películas en 3D, la conducción autónoma, el análisis de movimiento y actividad, y la interacción humano-robot. Según la dimensión espacial de los resultados de salida, las tareas principales de MHPE se dividen en dos categorías: estimación de pose 2D y estimación de pose 3D [3].

La estimación de pose humana 2D monocular, también conocida como detección de puntos clave 2D, tiene como objetivo localizar las coordenadas 2D de los puntos clave anatómicos humanos (articulaciones del cuerpo) a partir de imágenes. Debido a la falta de consenso en la comunidad científica sobre una taxonomía bien estructurada, no es útil profundizar en los pasos específicos de los diferentes métodos. Si bien algunos métodos comparten etapas similares, otros siguen enfoques diferentes en la estimación de la pose humana. De cualquier forma, podemos resumir el proceso en siete sencillos pasos cuya secuencia se representa en la *Figura 2*:

1. El método se clasifica según el uso de un modelo predefinido del cuerpo humano, lo cual determinará si es generativo (basado en modelo), discriminativo (libre de modelo) o una combinación de ambos enfoques (híbrido).
2. La información en 2D o 3D se utiliza como fuente de datos y/o medida de precisión para el método.

3. Se aplican técnicas de preprocesamiento, como la extracción del fondo de la imagen, la eliminación de ruido o la utilización de cajas delimitadoras (*bounding boxes*), para mejorar la calidad de los datos.
4. Se extraen y seleccionan características clave del cuerpo humano para su posterior análisis.
5. Se obtiene una pose inicial en 3D utilizando los datos y características recolectados.
6. Se aplican modelos y restricciones para descartar poses no realistas y refinar la estimación.
7. Se obtiene el modelo final de la pose humana, el cual representa la estimación más precisa y realista posible.

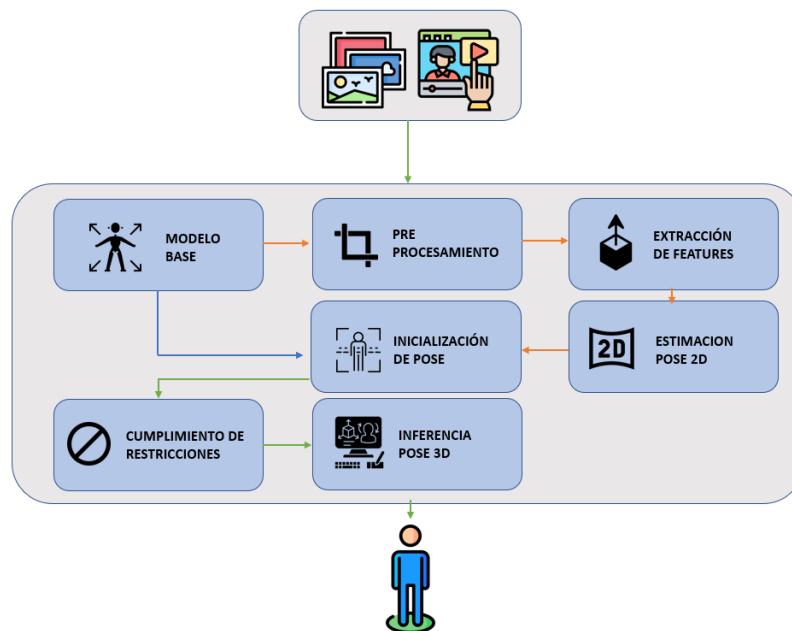


Figura 2: Pasos que se siguen para obtener la pose 3D dada una entrada de imágenes o video (Fuente: Elaboración propia basada en la Fig.3 de Sarafianos et al. (2016) [5])

Por otro lado, como se muestra en la *Figura 3*, desde el año 2014 hasta la actualidad, el número de artículos publicados en las principales conferencias (CVPR, ICCV, ECCV, etc.) y revistas (TPAMI, TIP, TOG, etc.) en el área de visión por computadora, multimedia y gráficos por computadora ha aumentado rápidamente. Los trabajos recientes se centran en el diseño y la optimización de redes, la interacción multitarea y la exploración de modelos corporales. Aunque se han logrado grandes éxitos en rendimiento y práctica, pocos trabajos han revisado de manera integral los algoritmos representativos o han realizado análisis perspicaces.

Desde la aparición de los modelos estadísticos del cuerpo, los investigadores los han utilizado para estimar la forma y la pose a partir de imágenes monoculares. Balan et al. [14] fueron

pioneros en estimar los parámetros de SCAPE [15] a partir de imágenes. SMPL [16] se impuso en el ámbito académico para la recuperación de la forma tridimensional del cuerpo. El mérito se debe a la naturaleza de código abierto de SMPL y a su comunidad en constante desarrollo: los métodos de adquisición de *ground-truth* [17], [18], conjuntos de datos con anotaciones extendidas de SMPL [19], [20], [21], y trabajos destacados [22], [23].

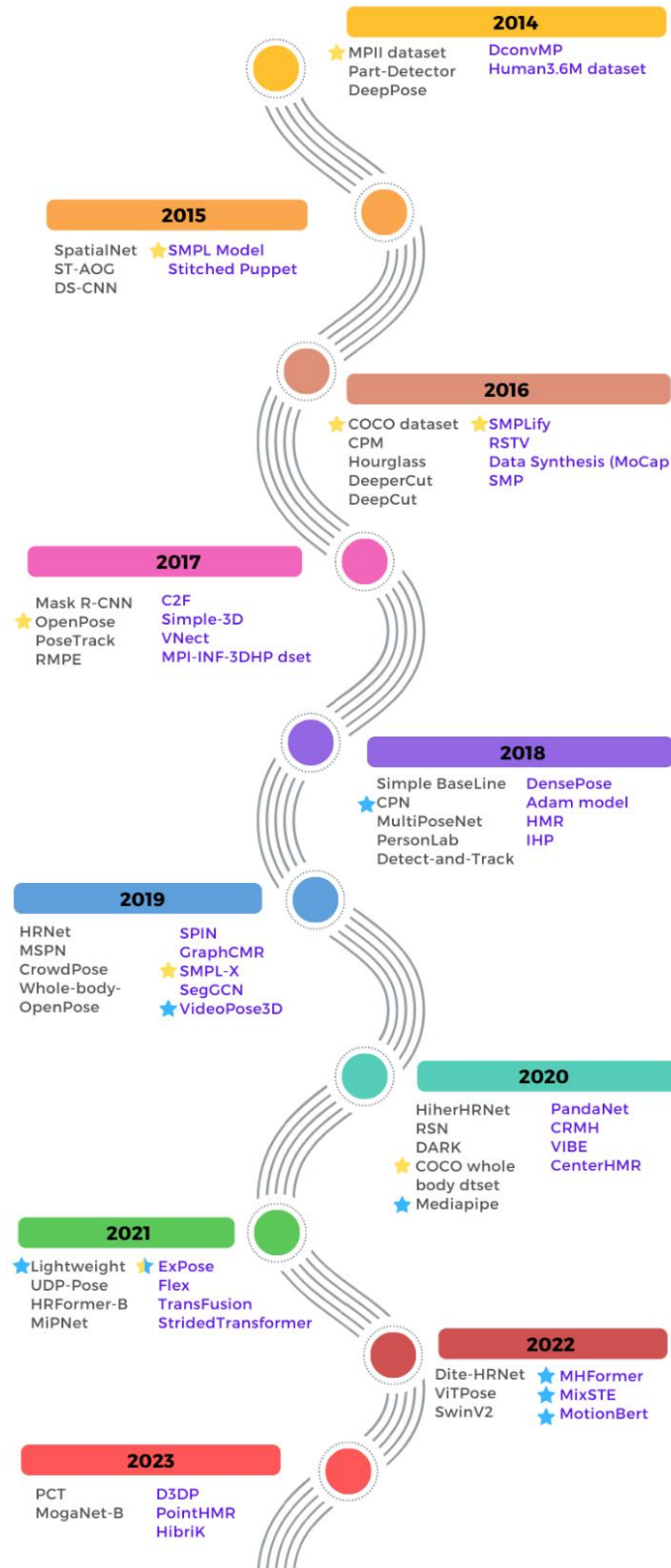


Figura 3: Hitos, avances en ideas o conjuntos de datos y los métodos más avanzados para la estimación de poses en 2D (en gris) y 3D (en azul). Las estrellas amarillas indican una especial importancia en el proyecto, las estrellas azules indican la actual presencia en la librería Bodyflow. (Elaboración propia)

En los últimos años, la mejora en las metodologías basadas en aprendizaje profundo (*Deep Learning*), la alta capacidad de cómputo, el acceso a bases de datos y librerías *open source*, han resultado en la creación y optimización de modelos de estimación de pose con un alto nivel de precisión.

En el estado del arte se pueden diferenciar aquellos métodos que estiman la postura en 2D, es decir, identifican las articulaciones del cuerpo (*keypoints*) usando imágenes, y aquellos que predicen la postura en 3D, mediante la inferencia de la profundidad a la que se encuentra la articulación. Este último método comúnmente se basa en la inferencia en 2D como un paso previo a la 3D. Cabe destacar que, a su vez, se pueden distinguir estos algoritmos mediante su capacidad de analizar únicamente a una persona que aparezca en la imagen (*single-person*) o más de una persona en paralelo (*multi-person*). En la Figura 4, se presentan algunos métodos representativos que se categorizan según los modelos humanos que utilizan centrándonos en los modelos de "*mesh recovery*" (en castellano, recuperación de malla).

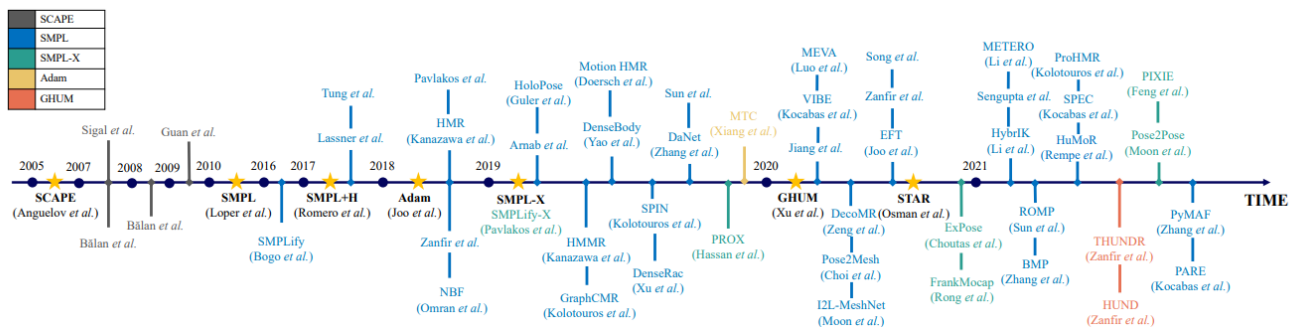


Figura 4: Representación cronológica de los modelos paramétricos más relevantes de humanos y los métodos de recuperación de malla 3D de humanos. [4]

Para abordar la estimación de poses en entornos con una sola persona o varias personas, se emplean enfoques denominados *top-down* y *bottom-up*, como se muestra en las Figuras 5 y 6. En el enfoque *top-down*, se identifican inicialmente a los individuos en la imagen mediante algoritmos de detección existentes, delimitándolos con cajas. A continuación, se predice la pose de cada individuo por separado.

En cambio, en el enfoque *bottom-up* se extraen todos los *keypoints* correspondientes a las articulaciones y luego se agrupan en individuos utilizando diversas técnicas, como la *clusterización*. Los algoritmos *top-down* suelen ofrecer una mayor precisión, pero requieren más recursos computacionales al utilizar dos detectores, mientras que los enfoques *bottom-up* son más rápidos, pero con una precisión ligeramente menor.

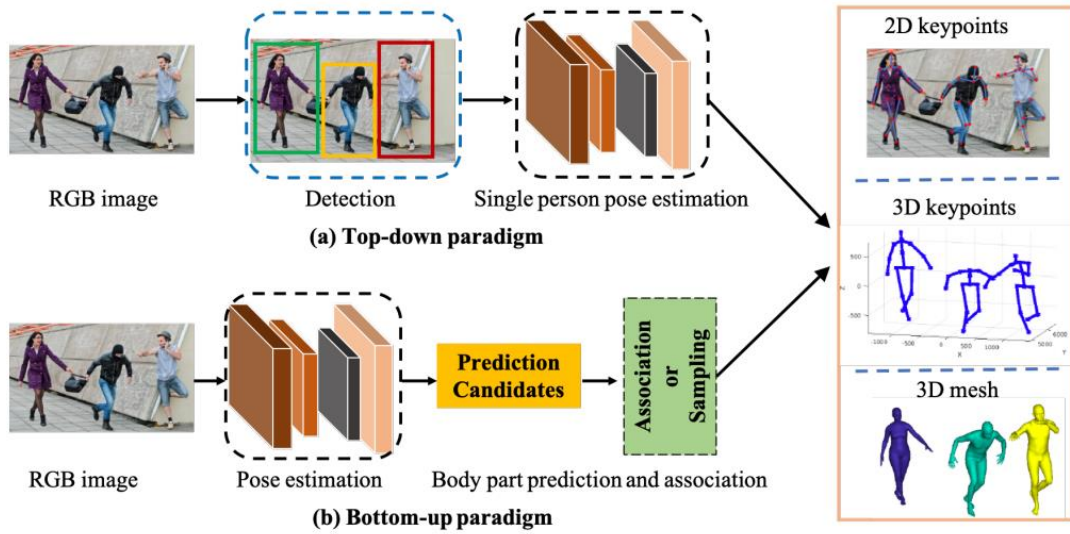


Figura 5: Paradigmas top-down y bottom-up de la detección de pose [2]

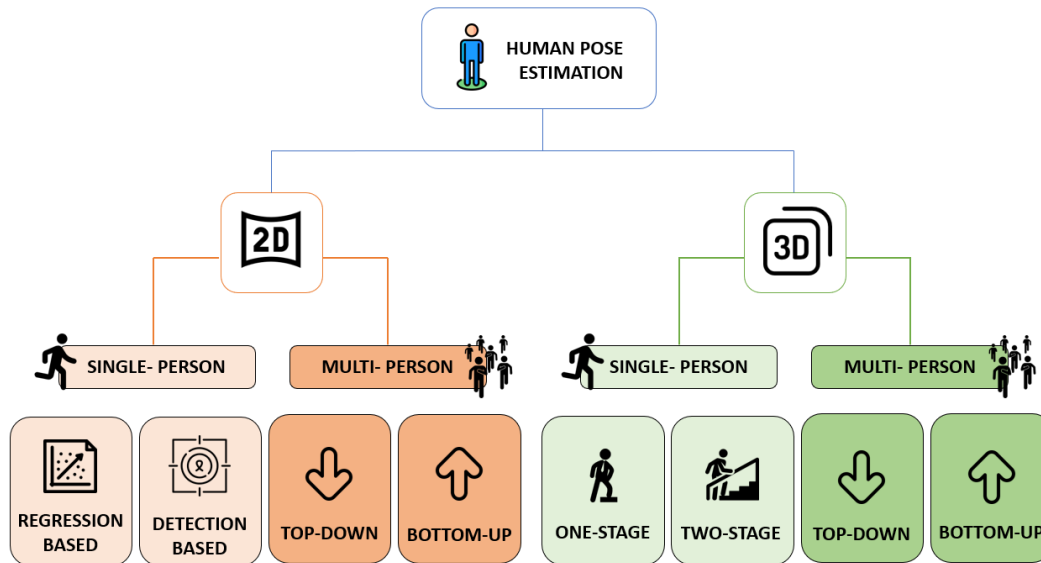


Figura 6: Tipos de estimación de pose humana (Elaboración propia)

Muchos son los acercamientos y las soluciones, aunque existen varias limitaciones importantes localizadas:

- Escasez de datos etiquetados en 3D: la falta de datos de entrenamiento en 3D dificulta el entrenamiento de redes neuronales y aumenta el riesgo de sobreajuste. Las bases de

datos existentes están grabadas en entornos controlados, lo que dificulta la generalización a situaciones reales en exteriores.

- b. Problemas de asociación de *keypoints*: cuando hay múltiples actores en una escena o se producen oclusiones, los modelos tienen dificultades para identificar los *keypoints* correspondientes a cada individuo. Este desafío está especialmente presente en enfoques *bottom-up*.
- c. Variaciones debidas a la proyección de la pose: la perspectiva de la cámara puede llevar a que diferentes poses 3D se mapeen a la misma pose 2D, lo que dificulta la inferencia precisa de la pose en 3D.
- d. Requisitos computacionales elevados: el cálculo de la pose humana en 3D requiere una gran cantidad de recursos computacionales, especialmente, cuando hay múltiples actores en la escena. Se plantean soluciones más rápidas y portables con redes neuronales más eficientes.
- e. Limitaciones de enfoques basados en mapas de calor: los sistemas de estimación de pose basados en mapas de calor pueden ofrecer resultados superiores, pero también presentan limitaciones, como la no linealidad del problema y el aumento del costo computacional. Los enfoques basados en regresión pueden ofrecer soluciones subóptimas, pero con una mayor velocidad y precisión.

2.1.2. Reconocimiento de actividades humanas con IA (HAR-Human Activity Recognition)

Pese a no adentrarme en exceso en este gran ámbito de detección de actividad, ya que el modelo integrado es de estimación de pose, sí que se utiliza para la obtención de resultados, por tanto, veamos de forma breve en qué consiste y en qué punto se encuentra.

El reconocimiento de la actividad humana (HAR) se enfoca en la identificación de actividades humanas mediante el uso de Inteligencia Artificial (IA) y la recopilación de datos provenientes de diversos dispositivos, como sensores ponibles, sensores inerciales de teléfonos móviles y cámaras de vídeo. El HAR tiene aplicaciones variadas, desde atención médica hasta vigilancia, y su adopción generalizada beneficia la seguridad y calidad de vida de las personas. En los últimos años, el aprendizaje profundo (DL-*Deep Learning*) ha revolucionado el campo del aprendizaje automático (ML) tradicional, mejorando el rendimiento en áreas como reconocimiento de imágenes, detección de objetos, reconocimiento de voz y procesamiento del lenguaje natural. El DL ha mejorado la solidez y rendimiento del HAR, permitiendo su implementación en una amplia gama de dispositivos.

El DL destaca por su capacidad de aprender características representativas de señales sin procesar sin requerir conocimiento experto y su capacidad de aproximar casi cualquier función con redes neuronales profundas. Si bien el DL se utiliza en diversas aplicaciones de reconocimiento de actividades humanas, este documento se centra específicamente en las actividades de la vida diaria (*ADL-Activities of Daily Living*) y en el reconocimiento de la marcha (*HGR-Human Gait Recognition*). El proceso de HAR generalmente consta de cuatro fases principales: adquisición de datos, preprocesamiento, entrenamiento del modelo y evaluación del desempeño.

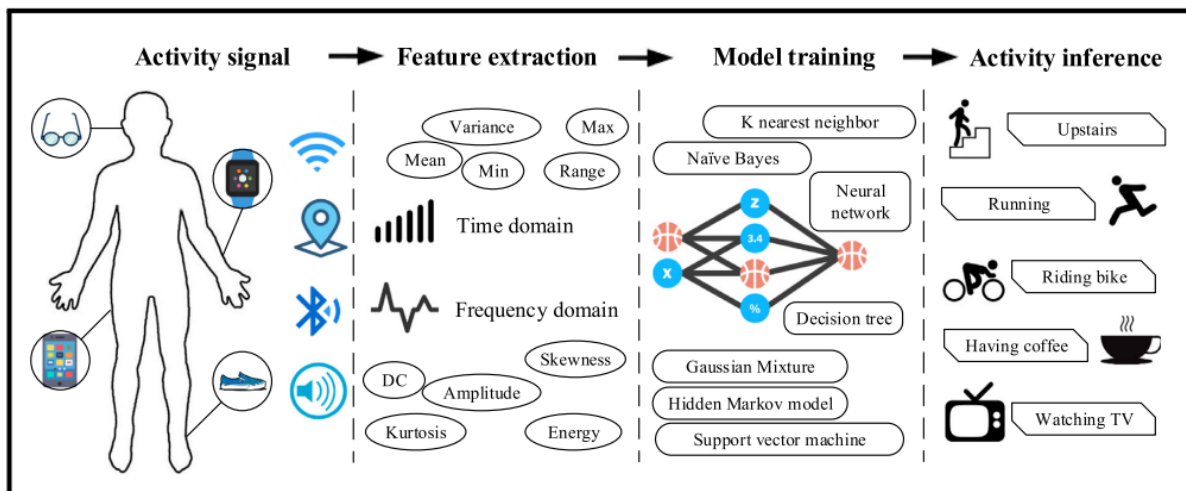


Figura 7: Una ilustración de reconocimiento de actividad basado en sensores utilizando enfoques convencionales de reconocimiento de patrones. [24]

En la primera fase, se elige un dispositivo de reconocimiento de actividad (HAR) según la aplicación. Por ejemplo, en una aplicación de vigilancia con muchas personas, las cámaras suelen ser el dispositivo HAR utilizado para la recopilación de datos. De manera similar, en una aplicación de monitoreo de la actividad diaria de una persona, uno o varios sensores suelen ser la fuente de adquisición de datos.

Los datos adquiridos pueden contener ruido y otras señales no deseadas, por lo que es importante contar con un sistema robusto de extracción de características para obtener mejores predicciones. Este proceso de adaptación de los datos y extracción de características relevantes se lleva a cabo en la segunda etapa. Una vez que los datos se han procesado, están listos para entrenar modelos de aprendizaje automático o aprendizaje profundo (fase 3 del proceso HAR). Después de la etapa 3, el modelo HAR se puede utilizar en la aplicación para realizar predicciones. La etapa 4 es la parte más desafiante, ya que el modelo se aplica a datos reales, cuyo comportamiento puede variar según factores físicos como la edad, la condición física y el enfoque utilizado para realizar una tarea.

Es importante destacar que el desarrollo de métodos HAR es un campo en constante evolución, como hemos mencionado anteriormente con la detección de pose, impulsado por avances en la tecnología de sensores, algoritmos de aprendizaje automático y hardware de procesamiento, a medida que se perfeccionan y se integran nuevas técnicas, el potencial de aplicaciones y sistemas basados en HAR sigue creciendo, lo que abre oportunidades emocionantes para la investigación y la innovación en el futuro cercano.

2.2. Antecedentes

El proyecto ha sido concebido y llevado a cabo en estrecha colaboración con el prestigioso Instituto Tecnológico de Aragón (ITAINNOVA), lo que ha sido un factor determinante para su éxito y relevancia en el campo de estudio. Este hecho ha fortalecido la credibilidad del proyecto, fomentando la innovación y abriendo oportunidades para futuras colaboraciones cercanas en diversas áreas tanto en el área clínica como deportiva.

Es fundamental tener en cuenta los antecedentes que han impulsado el desarrollo de este proyecto antes de adentrarnos en sus detalles. La detección de pose, especialmente en entornos tridimensionales, ha sido objeto de investigación y estudio durante un largo período de tiempo. A medida que avanzamos en la actualidad, presenciamos avances rápidos y continuos en este campo, con numerosos proyectos de investigación en marcha. Sin embargo, el verdadero desafío no radica únicamente en la investigación teórica, sino en la implementación práctica de estos avances. La aplicación efectiva de los métodos de detección de pose en entornos reales es un objetivo complejo que requiere superar obstáculos técnicos, optimizar algoritmos y adaptar las soluciones a las necesidades específicas de cada contexto.

Por lo tanto, este proyecto no solo se basa en los conocimientos y avances previos en el campo de la detección de pose, sino que también busca abordar el desafío de llevar estos avances a la práctica. La finalidad es lograr una implementación efectiva y funcional de las técnicas de detección de pose, que pueda ser aplicada en situaciones reales y proporcionar resultados precisos y confiables.

Un objetivo a largo plazo de la visión por computadora es comprender a los seres humanos y su comportamiento en escenarios cotidianos utilizando únicamente imágenes: ¿están felices o tristes?, ¿cómo interactúan entre ellos y con el mundo físico?, ¿cuáles son sus intenciones? Para responder a estas preguntas difíciles, primero necesitamos reconstruir de manera rápida y precisa su cuerpo, rostro y manos en 3D a partir de una única imagen RGB. Como resultado, la comunidad ha dividido el problema en partes, centrándose en gran medida en estimar el cuerpo principal o rompiendo el problema en piezas para estimar la expresión de la cara y la posición de las manos de manera separada [1].

Muchos métodos actuales estiman las articulaciones y características en 2D para el cuerpo, el rostro y las manos. Sin embargo, esto no es suficiente. Es la superficie de la piel la que describe aspectos importantes de los seres humanos, como su forma precisa en 3D, si están sonriendo, gesticulando o sosteniendo algo.

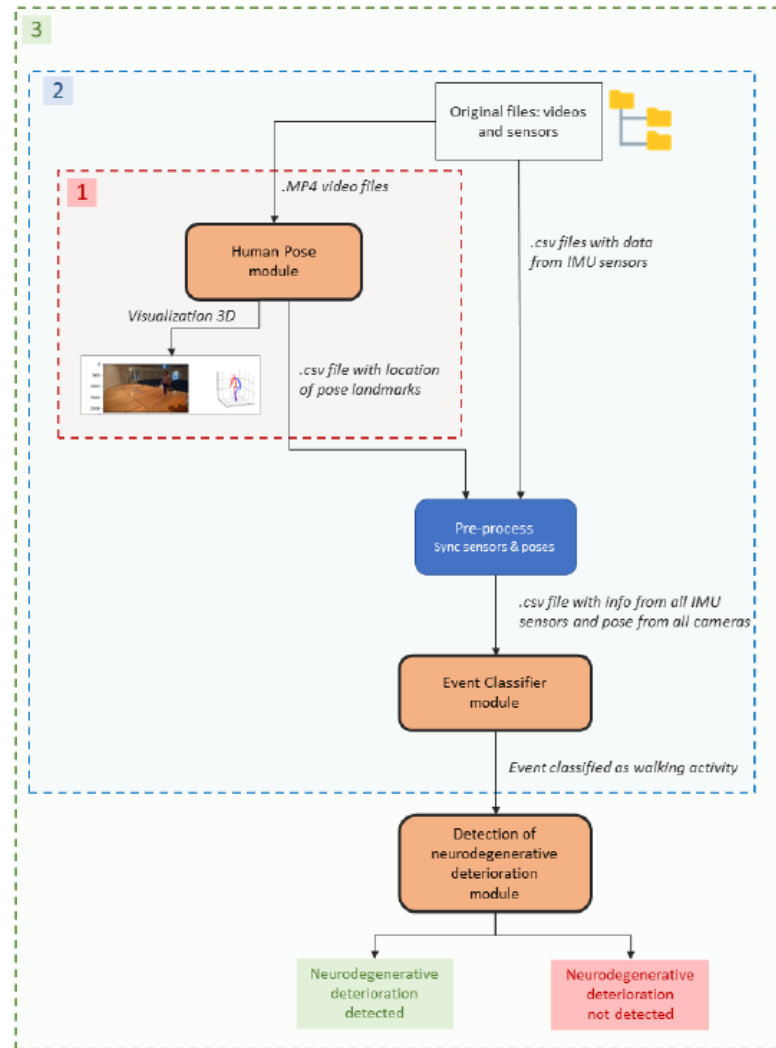


Figura 8: Estructura de la librería BodyFlow [44]

En el módulo de Estimación de Pose Humana, se emplean soluciones integradas que siguen una metodología de detección 2D, utilizando los puntos clave identificados para reconstruir una representación en 3D de dichos puntos y su forma correspondiente. Sin embargo, la integración de ExPose lleva el resultado a un nivel superior al proporcionar dos mejoras significativas. En primer lugar, se logra obtener una representación volumétrica en forma de malla, lo cual proporciona una descripción más detallada y precisa del cuerpo. En segundo lugar, se introduce la capacidad de estimar la pose de la cara y las manos en un único proceso, lo que simplifica el flujo de trabajo y brinda una visión más completa del sujeto en cuestión. Esta integración de ExPose en el módulo de Estimación de Pose Humana resulta en un *output* de mayor valor, al

combinar la representación 3D del cuerpo con la estimación facial y de las manos, en una solución integral y completa que se muestra a alto nivel en la *Figura 8*.

Como se ilustra en la *Figura 9*, la librería sigue una metodología de arriba hacia abajo (*top-down*) en el contexto de detección de pose. Esta metodología implica identificar inicialmente a la persona en la imagen, estimar su posición en 2D y luego utilizar esa información para predecir la posición en 3D. Este enfoque permite obtener resultados más precisos, pero depende en gran medida de un detector de pose 2D confiable. Una vez configurado el sistema según las preferencias del usuario, el programa se ejecuta y los detectores se inicializan en la memoria.

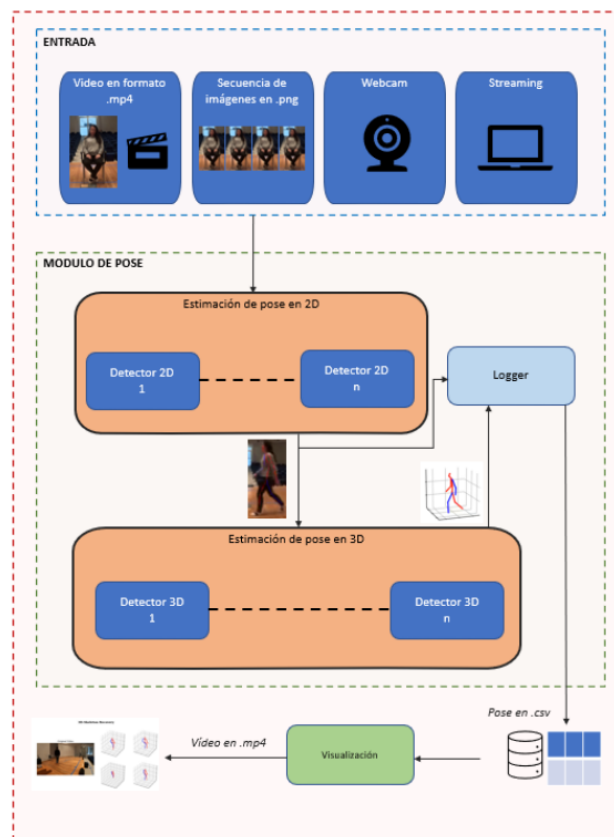


Figura 9: Módulo de Estimación de Pose (HPE) [44]

Después de inferir la pose de la persona en un fotograma, se almacena tanto la posición en 2D como en 3D. Una vez completado todo el proceso de predicción, se genera un archivo .csv que puede ser utilizado en otros contextos. Además, se incluye un módulo de visualización que permite verificar de manera visual que la pose ha sido inferida correctamente.

En el segundo módulo se pretenden utilizar datos procedentes de vídeos y de IMUs (*Inertial Measurement Units* o dispositivos electrónicos que miden e informan de la aceleración, la orientación, las velocidades angulares y otras fuerzas gravitatorias) [47] para clasificar la actividad que está desempeñando la persona. Partiendo del módulo de detección de pose, es

posible utilizar la pose de la persona que aparece en el vídeo en vez de utilizar el vídeo en crudo. En este caso, se probarán distintos modelos y se utilizará aquel con el que se obtengan los mejores resultados. En la *Figura 8* se puede observar dónde se enmarca este módulo, “*Human Activity Recognition*”, que tiene como entrada la información de la pose y los sensores.

En la *Figura 10* se muestra el funcionamiento general de este módulo. Se utilizan datos heterogéneos como entrada, los cuales son procesados para mejorar su significado. Se pueden aplicar técnicas como la extracción de estadísticas relevantes para el proceso de predicción. A continuación, se entrenan modelos con características destacadas, como el uso de convoluciones, jerarquías y otras técnicas avanzadas utilizadas en el estado del arte. Una vez que el modelo está entrenado, se realiza la predicción de la actividad que la persona está llevando a cabo en ese momento. Finalmente, el módulo de detección de deterioro cognitivo todavía se encuentra en desarrollo.

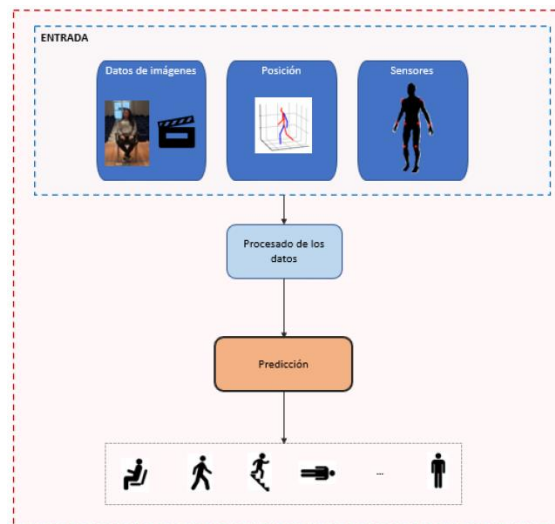


Figura 10: Módulo detección de eventos (HAR) [44]

3. Objetivos

En un principio, a 14 de octubre de 2022, el objetivo principal de este trabajo se centra en el desarrollo de una metodología para la estimación de pose 3D a partir de una representación reducida, de manera robusta y rápida, mediante algoritmos de ML. Para ello se propone estudiar los algoritmos de estimación de pose 3D más utilizados, determinando sus ventajas y desventajas. Además, se propone estudiar estos algoritmos de Inteligencia Artificial en diferentes situaciones ya que las poses que reconocerá la aplicación se definirán en función del contexto final de uso (como detección de caídas, actividades o eventos que le puedan suceder a una persona). La propuesta inicial se adjunta como *Anexo I*.

Una vez iniciado el proyecto, el objetivo inicial de la propuesta se concretó en los siguientes objetivos específicos:

- *Estudio y comparación de algoritmos de estimación de Pose 3D:* se investiga y analiza los algoritmos existentes para la estimación de pose 3D
- *Evaluación situacional de los algoritmos:* se evalúa el rendimiento de los algoritmos de estimación de pose 3D en diferentes situaciones o contextos de uso específicos. Esto puede incluir la detección de caídas, actividades o eventos que puedan ocurrir en una persona.
- *Estudio de la librería open source BodyFlow:* definición de las diferentes funcionalidades que están o estarán presentes en la herramienta ITAINNOVA para la detección de enfermedades neurodegenerativas y sarcopenia a través de la detección de eventos y el análisis de la pose.
- *Estudio sobre SMPL, SMPL-X, ExPose:* el estudio se centra en modelos y técnicas que permiten la síntesis y generación de animaciones de personajes humanos en 3D con alto nivel de detalle y realismo

El proyecto ha experimentado cambios menores a lo largo del tiempo para adaptarse a diferentes situaciones, planteándose las siguientes extensiones respecto al objetivo inicial:

- *Integración de un nuevo módulo "ExPose" en la librería de Bodyflow:* el objetivo principal es incorporar un módulo adicional de predicción 3d a la librería de Bodyflow. Este módulo proporcionará funcionalidades adicionales como la captura expresiva del cuerpo. La integración de este nuevo módulo se lleva a cabo de manera coordinada, garantizando la compatibilidad y la coherencia con el resto de los componentes de la librería.

- *Visualización propia:* este objetivo se centra en desarrollar una función de visualización propia dentro de la plataforma para conseguir visualizar la información obtenida del modelo integrado.
- *Actualización de la función de sincronización:* modificación de la función ya existente para su actividad con las nuevas funcionalidades
- *Utilización del módulo de HAR para la obtención de métricas:* se utiliza el módulo de HAR para obtener métricas de reconocimiento de actividades humanas con ExPose y se comparan con las métricas obtenidas por ITAINNOVA.

4. Metodología

El proyecto está planteado para utilizar una metodología ágil ya que se caracteriza por su enfoque flexible, lo que lo hace adecuado para abordar proyectos complejos y en constante cambio. Al elegir una metodología ágil, se buscaba mejorar la capacidad de adaptarse a los requisitos cambiantes del cliente y mantener una comunicación fluida con todas las partes interesadas.

Además, la metodología ágil fomenta la entrega temprana y frecuente de incrementos del producto, lo que me permite obtener retroalimentación temprana y realizar ajustes en función de las necesidades reales del usuario. Esto ayuda a reducir los riesgos asociados con el desarrollo de un producto y garantiza que el equipo se mantenga enfocado en entregar valor en cada etapa del proyecto.

La transparencia y la visibilidad son pilares fundamentales de las metodologías ágiles, lo que significa que todos los miembros del equipo pueden ver el progreso y el estado del proyecto en tiempo real. Esto facilita la toma de decisiones informadas y permite identificar y abordar problemas rápidamente.

4.1. Bases del proyecto

En el caso específico de este proyecto, la metodología ágil también permite dividir el trabajo en iteraciones o *sprints* más pequeños y manejables, lo que facilita el seguimiento del progreso y la identificación temprana de posibles desviaciones. Además, se ha podido priorizar el trabajo según el valor que aporta al producto final, lo que ayuda a asegurar que las características más importantes se implementen primero.

Primero se descartaron ciertas metodologías como XP debido al tamaño y complejidad del proyecto, la necesidad de adaptabilidad y la naturaleza específica del equipo o Lean porque no se ajustaba a las características del proyecto ya que se centra en la eliminación de desperdicios y la mejora continua, lo cual puede ser adecuado para proyectos más pequeños o procesos de producción, pero podría no ser la mejor opción para un proyecto de mayor envergadura que requiere una estructura más definida y planificación específica.

Después se planteó la posibilidad de utilizar DSDM (*Dynamic Systems Development Method*) pero también fue descartado debido a que esta metodología se centra en proyectos de desarrollo de software. Como el proyecto no estaba estrictamente enfocado en desarrollo de software, DSDM no era la opción más adecuada.

Tras descartar estas metodologías, se consideran otras opciones ágiles, la elección entre Scrum, Kanban y Scrumban depende de las necesidades y circunstancias específicas de un equipo u organización. A continuación, discutiremos brevemente cada enfoque para comprender por qué de la elección de Scrumban en lugar de Scrum o Kanban [26]:

Por un lado, Scrum es un marco de trabajo ágil ampliamente utilizado que se basa en roles definidos, ceremonias e iteraciones acotadas en el tiempo llamadas *sprints*. Es adecuado para proyectos con requisitos claros y estables, donde los comentarios y la colaboración frecuente son esenciales. Scrum proporciona un enfoque estructurado para el desarrollo de software y es ideal para equipos que buscan un proceso bien definido.

Por otro lado, Kanban es una metodología ágil que se centra en la entrega continua y en visualizar el trabajo en un tablero Kanban. Es adecuado para equipos con requisitos que cambian con frecuencia o para equipos de soporte donde las prioridades pueden cambiar rápidamente. Kanban permite a los equipos adaptarse a los cambios de manera más efectiva y proporciona flexibilidad en la gestión del flujo de trabajo.

4.2. ¿Por qué Scrumban?

Ahora podemos entender mejor Scrumban que no deja de ser una combinación de Scrum y Kanban que combina lo mejor de ambos enfoques. Scrumban agrega principios de Kanban, como el trabajo basado en *pull* y el flujo continuo, a Scrum, lo que lo hace más adaptable y menos prescriptivo [27]:

- a. Menor rigidez: En Scrum algunos aspectos son demasiado rígidos o difíciles de seguir, Scrumban permite al equipo conservar elementos de Scrum mientras adopta prácticas más flexibles de Kanban.
- b. Demanda Variable: Si el trabajo del equipo tiene una demanda variable y prioridades que cambian con frecuencia, el enfoque basado en flujo de Scrumban puede proporcionar una mejor visibilidad y adaptabilidad.
- c. Entrega Continua: Para equipos enfocados en la entrega continua y que desean minimizar el desperdicio y reducir retrasos, el enfoque basado en pull de Scrumban es beneficioso.
- d. Mejora Continua: Scrumban fomenta la mejora continua, lo que permite a los equipos evolucionar sus procesos gradualmente y adaptarse a las necesidades cambiantes con el tiempo.
- e. Tipos Mixtos de Proyectos: Si un equipo maneja una mezcla de proyectos, algunos de los cuales podrían beneficiarse de la estructura de Scrum y otros de la flexibilidad de Kanban, Scrumban puede encontrar un equilibrio entre ambos.

En resumen, Scrumban es una opción pragmática cuando un equipo desea conservar la estructura de Scrum mientras adopta los principios de flujo y adaptabilidad de Kanban. Además, es de mencionar que este proyecto es de tamaño modesto y tiene un propósito académico, con solo un grupo pequeño de personas involucradas en él. El objetivo principal es crear una solución sólida, pero su enfoque principal es permitir que el estudiante que lo desarrolla aprenda mediante

la práctica. Además, se valora mucho el tiempo y se busca evitar dedicar más tiempo a la metodología que al trabajo real. A continuación, se describen los conceptos clave de ScrumBan y su aplicación en el proyecto:

4.2.1. Scrum en Scrumban

- Roles: ScrumBan mantiene los roles básicos de Scrum, que incluyen
 - Scrum Master: Rafael del Hoyo Alonso. Responsable de asegurarse de que el equipo siga las prácticas ágiles y Scrum de manera efectiva. También ayuda a eliminar obstáculos y facilita el trabajo del equipo.
 - Product Owner: ITAINNOVA. En el caso de nuestro departamento es Alberto Capella el responsable de gestionar el backlog del producto y de tomar decisiones sobre las características y prioridades del producto. Se asegura de que el equipo desarrolle las funcionalidades más valiosas para el cliente.
 - Equipo de Desarrollo: Marcos Marina Castelló
 - Sprints: Aunque ScrumBan se basa en Kanban, se utilizan sprints para establecer intervalos de tiempo para la entrega y revisión de funcionalidades. Sin embargo, los sprints en ScrumBan tienen una duración más flexible que los sprints tradicionales de Scrum. En mi caso el trabajo se ha desarrollado en 5 sprints de duración y esfuerzo (en horas) variables como se muestra en el apartado de *Resultados*.

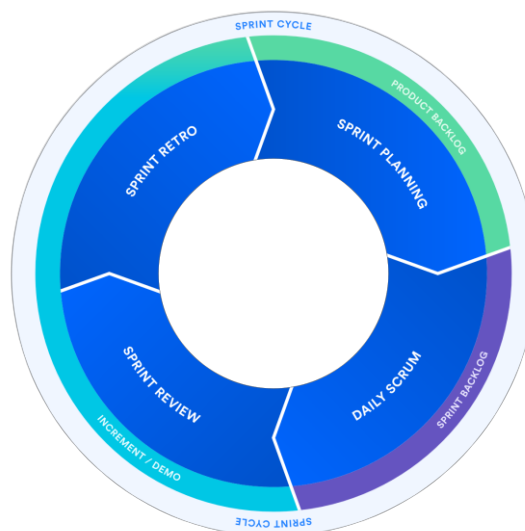


Figura 11: Gráfico ciclo de sprint [28]

- Reuniones: Se mantienen las reuniones tradicionales de Scrum, como la Planificación del Sprint, la Revisión del Sprint y la Retrospectiva del Sprint. Sin embargo, el enfoque en ScrumBan es permitir la flexibilidad en la planificación y el tiempo dedicado a estas reuniones.

4.2.2. Kanban en Scrumban

- Tablero Kanban: ScrumBan utiliza un tablero Kanban para visualizar el flujo de trabajo del equipo. El tablero muestra las tareas en diferentes etapas, como "Por hacer", "En progreso" y "Terminado". Cada tarea está representada por una tarjeta y se mueve a lo largo del tablero a medida que progresa en el proceso.
- Límites de trabajo en progreso (WIP): Al igual que en Kanban, ScrumBan utiliza límites WIP para controlar la cantidad de trabajo que puede estar en progreso a la vez. Esto ayuda a evitar la sobrecarga manteniendo un flujo de trabajo constante. Claramente se puede observar en el proceso de trabajo escrito.

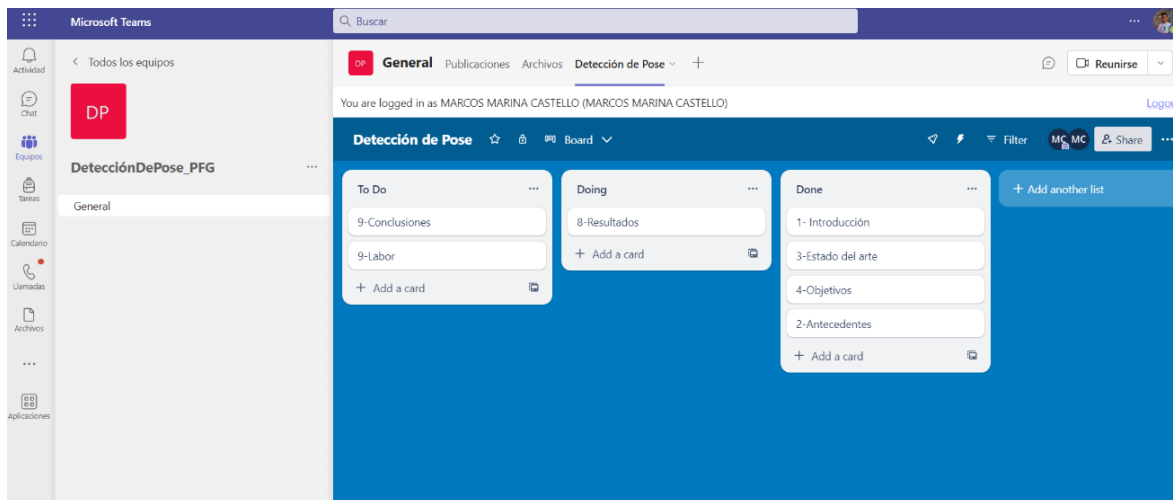


Figura 12: Imagen de Trello integrado en Teams para el desarrollo del trabajo escrito

- Enfoque en flujo continuo: A diferencia de Scrum, donde las funcionalidades se planifican en sprints, ScrumBan se centra en el flujo continuo de trabajo. Las tareas se toman y completan según sea necesario, sin una planificación estricta a largo plazo. Esto es evidente en el tablero de trabajo integrado en GitLab.

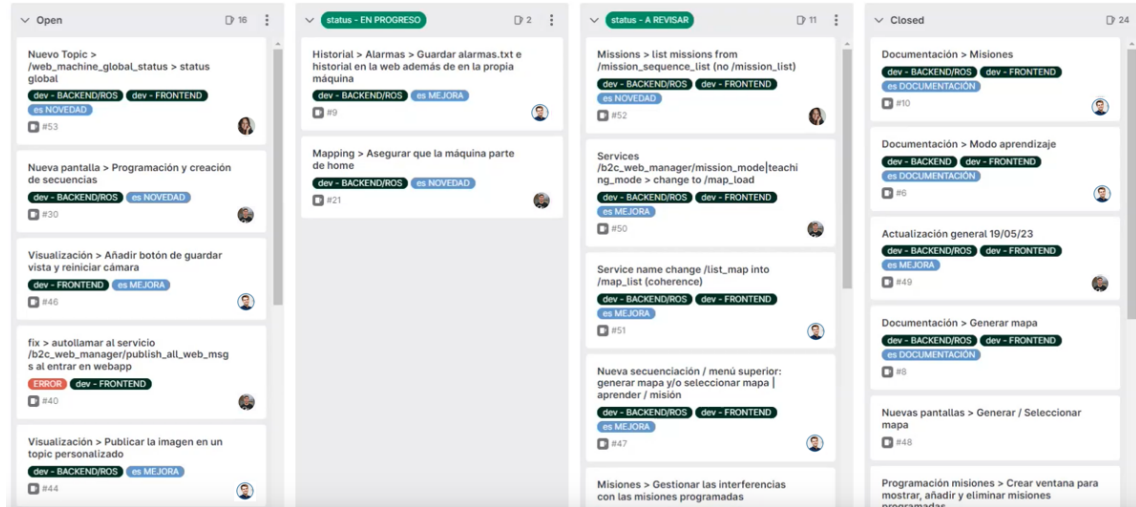


Figura 13: Ejemplo Tablero Kanban con enfoque en el desarrollo continuo GitLab.

Como se ha podido advertir, una de las principales características de ScrumBan es su flexibilidad. Permite a los equipos adaptarse y ajustar su proceso según sea necesario. Esto es útil cuando los requisitos del proyecto cambian con frecuencia o cuando el equipo necesita responder rápidamente a los cambios como ha sido necesario a lo largo del proyecto realizado.

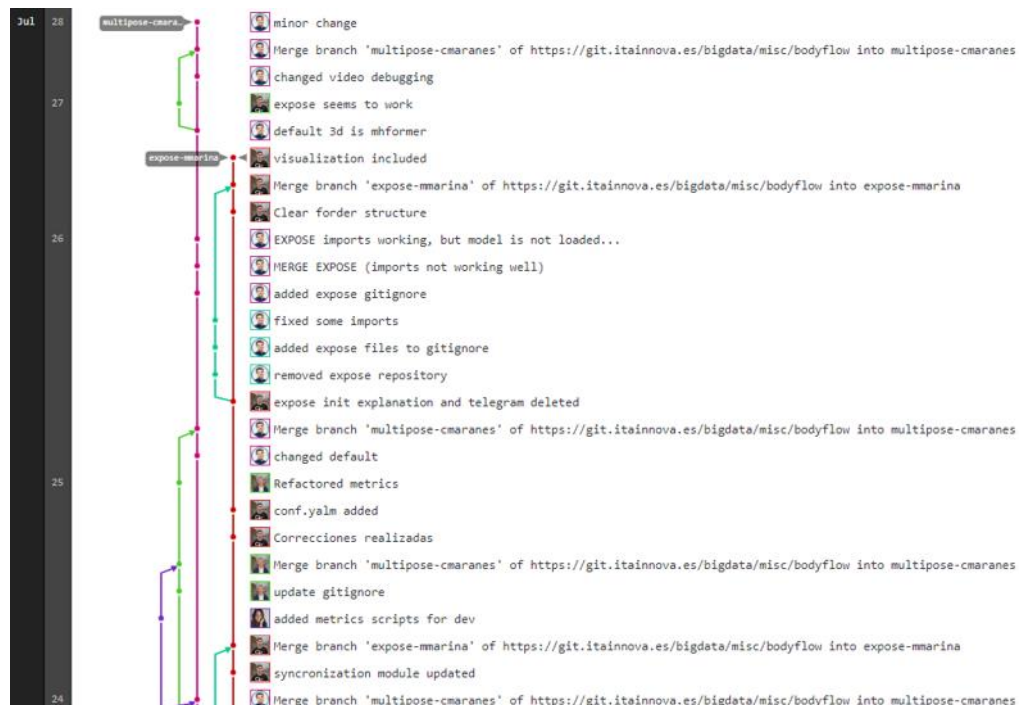


Figura 14: Captura del gráfico del repositorio en la etapa de integración

4.3. Gestión de versiones

El control de versiones del código fuente de BodyFlow se ha construido utilizando el sistema de control de versiones Git, el cual fue creado por Linus Torvalds y la comunidad de desarrolladores de Linux.

El objetivo principal de esta comunidad era desarrollar un sistema rápido y sencillo que pudiera manejar múltiples líneas de desarrollo en paralelo, ser completamente distribuido y tener un rendimiento óptimo para proyectos de gran envergadura, como el kernel de Linux. Git se lanzó en 2005 y ha revolucionado el mundo del desarrollo de software, ofreciendo una alternativa mucho más eficiente y efectiva para el control de versiones. Desde entonces, Git se ha convertido en una herramienta esencial en el desarrollo de software, permitiendo a los equipos de trabajo gestionar y colaborar en proyectos de manera organizada y eficiente.

4.3.1. GitLab y GitHub

El código se aloja en GitLab, una plataforma de gestión de repositorios de código fuente basada en Git. Es una herramienta de desarrollo colaborativa que permite a los equipos de trabajo almacenar, administrar y colaborar en proyectos de software de manera eficiente. GitLab proporciona una interfaz web que facilita la creación, clonación y administración de repositorios Git, así como funciones de seguimiento de problemas (*issue tracking*), integración continua, pruebas automatizadas, despliegue automático y mucho más.

Además de la versión en la nube que ofrece GitLab.com, también existe la opción de instalar GitLab en servidores locales o en la nube privada de una organización, lo que permite un mayor control y personalización.

Adicionalmente, ITAINNOVA cuenta con una cuenta de GitHub donde se hace público el código para facilitar su acceso y colaboración con la comunidad.

4.4. Estrategias de comunicación

La comunicación tanto con el Scrum Master como con el resto de los desarrolladores se lleva a cabo mediante la plataforma Microsoft Teams o de manera presencial en las instalaciones de ITAINNOVA.

Una vez entrado en vigor el contrato con la institución, como es lógico en el desarrollo de un proyecto de esta envergadura se realizan tanto reuniones diarias, en caso de tener dudas o necesidad de coordinación, como reuniones semanales a nivel de proyecto. Estas reuniones son fundamentales para mantener una comunicación efectiva entre los miembros del equipo, revisar el progreso del proyecto, identificar posibles problemas o desafíos, y tomar decisiones importantes. Además de las reuniones semanales a nivel de proyecto, que permiten evaluar el estado general del proyecto y asegurarse de que se están cumpliendo los objetivos y plazos

establecidos, se realizan reuniones a nivel de departamento de manera bisemanal, donde, se abordan temas más amplios y estratégicos relacionados con el área de trabajo del equipo.

Por otro lado, la comunicación con los tutores académicos se realiza a través de la cuenta de Teams de la universidad o en la propia universidad, dependiendo de las preferencias y disponibilidad de ambas partes.



5. Estudio, diseño e implementación

En este capítulo se presentarán las distintas tareas realizadas en torno al proyecto siguiendo la metodología Scrum, lo que implica trabajar en *sprints*. Al finalizar cada *sprint*, se llevará a cabo una revisión del tablero Kanban con el Scrum Master.

Tras la realización de las primeras reuniones se definen los objetivos que hemos mencionado anteriormente. Además se tiene en cuenta los cometidos propios de este proyecto como la introducción a esta área de conocimiento o la escritura de la memoria. Estos objetivos y cometidos se dividen en tareas que directamente se proyectan en el tablero *sprint a sprint*.

5.1. *Sprint 1: Estimación de Pose, ExPose, Bodyflow*

El proyecto requiere de unos conocimientos en inteligencia artificial y visión que se salen del currículo académico del grado, por lo que es necesario un proceso de documentación y aprendizaje. Este proceso de adquisición de conocimientos es fundamental para garantizar el desarrollo exitoso del proyecto. Además, este enfoque de aprendizaje continuo contribuirá a mantenerse actualizado con las últimas tendencias y avances tecnológicos en esta área de conocimiento, lo que resulta esencial en proyectos de alta complejidad y envergadura como este.

5.1.1. *Configuración del tablero Kanban*

En este sprint se va a trabajar sobre los objetivos *Estudio y comparación de algoritmos de estimación de Pose 3D* y *Estudio sobre SMPL, SMPL-X, ExPose*. El tiempo estimado es de 70 horas.

<u>Tarea 1</u>	Aprendizaje en conceptos básicos y funcionamiento red neuronal
<u>Tarea 2</u>	Estudio SMPL
<u>Tarea 3</u>	Estudio SMPL-X
<u>Tarea 4</u>	Estudio ExPose
<u>Tarea 5</u>	Comprensión utilidad y usabilidad Bodyflow

5.1.2. *Tarea 1- Aprendizaje en conceptos básicos y funcionamiento red neuronal*

Esta tarea se centra en adquirir conocimientos fundamentales sobre las redes neuronales y recapitular los aprendidos en el grado. Se estudian conceptos clave que permitirán comprender cómo funcionan estas estructuras y cómo se utilizan en diversos contextos [32] [33] [34]. A continuación, se describe el proceso de esta tarea:

- Introducción a las redes neuronales: Se inicia con una introducción general sobre qué son las redes neuronales y su relevancia en el campo del aprendizaje automático y la inteligencia artificial.

- Arquitectura de una red neuronal: Se estudia la estructura básica de una red neuronal, que incluye capas de neuronas y cómo estas capas se conectan entre sí. Se describen los diferentes tipos de capas, como capas de entrada, capas ocultas y capas de salida.
- Funciones de activación: Se analizan las funciones de activación utilizadas en las neuronas para introducir no linealidad en la red. Se exploran algunas funciones comunes, como la función ReLU (*Rectified Linear Unit*) y la función sigmoide.
- El proceso de entrenamiento: Se explican los pasos involucrados en el entrenamiento de una red neuronal utilizando algoritmos de optimización como el descenso del gradiente. Se comprende cómo la red ajusta sus pesos y parámetros para mejorar su rendimiento en una tarea específica.
- Retropropagación (*Backpropagation*): Se estudia el algoritmo de retropropagación, que es fundamental para el entrenamiento de redes neuronales. Se detalla cómo se calculan los gradientes y se propagan hacia atrás para ajustar los pesos de las conexiones.
- Funcionamiento en problemas de clasificación y regresión: Se explican los enfoques de clasificación y regresión en el contexto de redes neuronales. Se observan y replican ejemplos prácticos de cómo las redes pueden abordar tareas como reconocimiento de imágenes, clasificación de texto o predicción numérica.
- Overfitting y Regularización: Se aborda el problema de *overfitting* y cómo se puede aplicar la regularización para evitar este fenómeno en el entrenamiento de redes neuronales.
- Aplicaciones y tendencias: Se exploran diversas aplicaciones de redes neuronales en la vida real, como el procesamiento de imágenes, el procesamiento del lenguaje natural, la visión por computadora, entre otras. Además, se revisan las tendencias actuales en el campo del aprendizaje profundo y las redes neuronales.

5.1.3. Tarea 2- Estudio SMPL

En esta tarea, se lleva a cabo un estudio en profundidad sobre SMPL (*Skinned Multi-Person Linear Model*). Se investiga cómo funciona este modelo para estimar la pose 3D de personas, su arquitectura, las matemáticas detrás de él y los datos que utiliza [16][35]. Además, analizan trabajos académicos y artículos relevantes sobre SMPL para entender sus ventajas y limitaciones.

Por otro lado, se ahonda en SMPLify [17][36], un método para estimar automáticamente la pose 3D del cuerpo humano y su forma 3D a partir de imágenes 2D no restringidas. Utiliza un enfoque de optimización iterativa para ajustar el modelo SMPL a las ubicaciones de las articulaciones 2D detectadas en la imagen. SMPLify utiliza un enfoque donde primero predice las ubicaciones de

las articulaciones 2D mediante un modelo basado en redes neuronales (*DeepCut*) y luego ajusta el modelo 3D SMPL a esas ubicaciones para obtener la pose y forma 3D estimadas.

5.1.4. Tarea 3- Estudio SMPL-X

En esta tarea se continúa en la misma línea de trabajo, se realiza el estudio sobre SMPL-X, centrándose en investigar y profundizar en los aspectos clave del modelo mejorado de *Skinned Multi-Person Linear Model* [22]. En esta tarea además se profundiza en la licencia y en el código [37] para realizar pruebas con la demo que ponen a disposición de la comunidad.

Para facilitar el análisis de acciones humanas, interacciones y emociones, se calcula un modelo 3D de la pose del cuerpo humano, la pose de las manos y la expresión facial a partir de una única imagen monocromática. Para lograr esto, utilizan miles de escaneos 3D para entrenar un nuevo modelo 3D unificado del cuerpo humano, SMPL-X, que amplía SMPL con manos completamente articuladas y una cara expresiva. Aprender a regresar los parámetros de SMPL-X directamente desde imágenes es desafiante sin imágenes emparejadas o *Ground Truth*.

En consecuencia, se sigue el enfoque de SMPLify, que estima características 2D y luego optimiza los parámetros del modelo para ajustarse a las características. Por supuesto se mejora SMPLify de varias formas significativas:

- Se detectan características 2D correspondientes a la cara, las manos y los pies y se ajusta el modelo completo de SMPL-X a estas.
- Se entrena una nueva red neuronal para la pose previa utilizando un gran conjunto de datos de captura de movimiento (*MoCap*).
- Se define una nueva penalización de interpenetración asegurándose de que todas las partes del cuerpo se coloquen de manera coherente y no se intersecten en posiciones anatómicamente incorrectas.
- Se detecta automáticamente el género y los modelos corporales apropiados (masculino, femenino o neutro).
- Se implementa en *PyTorch* logrando una aceleración de más de 8 veces sobre *Chumpy*.

SMPL-X es un gran paso en la capacidad de capturar y representar automáticamente la postura y la expresión de personas que muestran emociones a partir de datos RGB monoculares.

5.1.5. Tarea 4- Estudio ExPose

En esta etapa final de la investigación, se realiza un análisis exhaustivo de ExPose, un método innovador que permite capturar de manera veloz y precisa el cuerpo humano, la cara y las manos en 3D a partir de imágenes RGB. Además de profundizar en su funcionamiento, se investiga su licencia y código, así como las diversas aplicaciones prácticas que ofrece en distintos campos.

Algunos enfoques recientes reconstruyen seres humanos expresivos completos en 3D a partir de imágenes utilizando modelos de cuerpo en 3D que incluyen la cara y las manos. Sin embargo, estos métodos se basan en la optimización, lo que los hace lentos, susceptibles a óptimos locales y requieren puntos clave 2D como entrada.

Para abordar estas limitaciones se presenta ExPose (*Expressive POse and Shape rEgression*), que regresa directamente el cuerpo, la cara y las manos, en formato SMPL-X, a partir de una imagen RGB. Esto no debemos pasar por alto que es un problema complejo debido a la alta dimensionalidad del cuerpo y la falta de datos de entrenamiento expresivos. Además, las manos y la cara son mucho más pequeñas que el cuerpo, ocupando muy pocos píxeles de imagen. Esto hace que la estimación de manos y cara sea difícil cuando se reducen las imágenes del cuerpo para las redes neuronales.

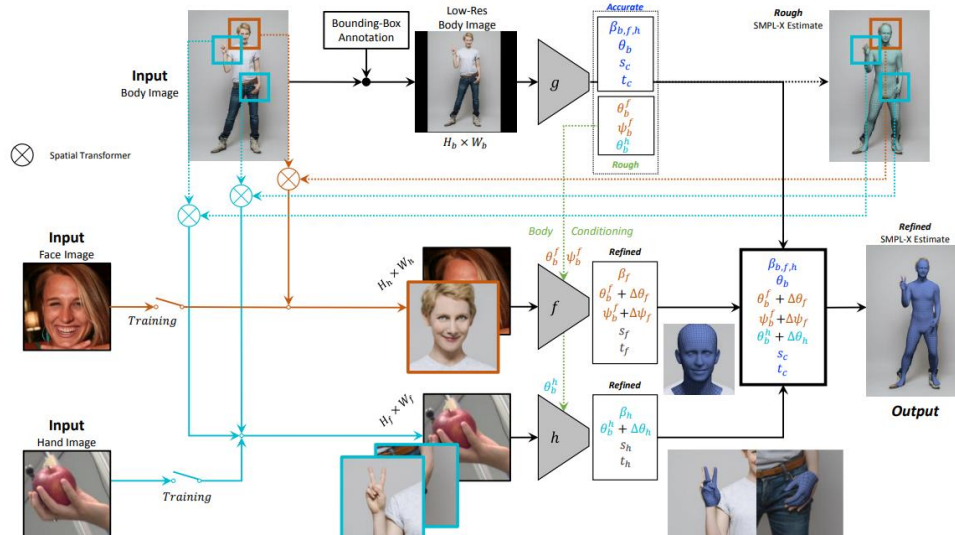


Figura 15: Se extrae una imagen del cuerpo utilizando un cuadro delimitador de la imagen completa y se ingresa a una red neuronal que predice la pose del cuerpo, manos y cara, así como la forma, expresión, escala de la cámara y traslación. Luego, se extraen imágenes específicas para la cara y las manos y se ingresan a subredes dedicadas para obtener estimaciones precisas [41].

Con ExPose se realizan tres contribuciones principales:

- Teniendo en cuenta la falta de datos de entrenamiento, se crea un conjunto de datos de ajustes de SMPL-X en imágenes del mundo real.
- Se observa que la estimación del cuerpo localiza razonablemente bien la cara y las manos de esta forma se introduce una atención guiada (*body-driven attention*) para las regiones de la cara y las manos en la imagen original extrayendo recortes de mayor resolución que alimentan los módulos de refinamiento dedicados.

- Estos módulos aprovechan el conocimiento específico de partes de conjuntos de datos existentes solo para la cara y las manos.

5.1.6. Tarea 5- Comprensión utilidad y usabilidad Bodyflow

En esta tarea, se realiza un análisis de Bodyflow, con el objetivo de comprender su utilidad y usabilidad en el contexto del proyecto o investigación actual. Bodyflow es una herramienta que permite estimar y capturar la pose 3D y las expresiones faciales de personas a partir de imágenes.

Primero se estudia la documentación oficial de Bodyflow para comprender su propósito, funcionalidades y capacidades analizando la información proporcionada por los desarrolladores y otros usuarios. Después se exploran las distintas funcionalidades y características que ofrece en la estimación de pose 3D y expresiones faciales examinando a su vez cómo se integra en el flujo de trabajo del proyecto. Finalmente se evalúa el rendimiento de Bodyflow en términos de precisión y eficiencia en la estimación de pose y expresiones faciales utilizando los datos generados por el equipo de ITAINNOVA.

Por otro lado, en esta tarea se está analizando la usabilidad de Bodyflow, evaluando su facilidad de uso y su adecuación a las necesidades y requerimientos del equipo de trabajo. Además, se examinan los requisitos necesarios para su correcto funcionamiento. También se identifican las ventajas y beneficios que Bodyflow aporta al proyecto, así como las posibles limitaciones o situaciones en las que podría no ser la mejor opción para la integración de un nuevo módulo por parte de un estudiante o investigador.

5.1.7. Evaluación del sprint

Durante este primer sprint, no se generó código propio ni se produjo un entregable final. Sin embargo, el Scrum Master ha aprobado el proceso de investigación en el que ha estado involucrado y ha participado como espectador en las demostraciones y explicaciones realizadas semanal o mensualmente. Aunque se dedicaron 72 horas en lugar de las 70 horas programadas, todas las tareas planificadas se llevaron a cabo con éxito.

5.2. *Sprint 2: Preparación del entorno y análisis del estado actual*

5.2.1. *Configuración tablero Kanban*

En este sprint se va a trabajar sobre los objetivos *Estudio de la librería open source BodyFlow y Estudio sobre SMPL, SMPL-X, ExPose*. El tiempo estimado es de 60 horas.

<u>Tarea 1</u>	Instalación y configuración Visual Studio Code – servidores ITAINNOVA
<u>Tarea 2</u>	Creación de ambientes de Conda
<u>Tarea 3</u>	Elaboración del Diagrama de clases para Bodyflow
<u>Tarea 4</u>	Elaboración del Diagrama de clases para ExPose
<u>Tarea 5</u>	Analizar ExPose para comprender su funcionamiento y exponerlo
<u>Tarea 6</u>	Realización de una Demo de ExPose

5.2.2. *Tarea 1- Instalación y configuración Visual Studio Code – servidores ITAINNOVA*

Pese a estar familiarizado con Visual Studio Code, un popular editor de código desarrollado por Microsoft que destaca por su versatilidad y extensibilidad en diferentes lenguajes de programación, se establece una reunión con Ana Hernández, Técnica I+D+I en Big Data de ITAINNOVA, donde se recibe una breve introducción al entorno de trabajo.

- **Servidores de ITAINNOVA:** Los servidores de ITAINNOVA son sistemas de alto rendimiento especialmente configurados para llevar a cabo aplicaciones y tareas que requieren un rendimiento superior. Estos servidores están configurados para soportar cargas de trabajo intensivas, como tareas de análisis de datos, procesamiento de imágenes, ejecución de algoritmos de inteligencia artificial y otras actividades computacionales complejas.
- **Gestión de *Environments*:** En el contexto de desarrollo, los *environments* o ambientes son entornos virtuales que permiten a los desarrolladores trabajar con diferentes versiones de lenguajes y bibliotecas de software. ITAINNOVA utiliza herramientas como Anaconda o Virtualenv para crear y gestionar entornos aislados, lo que facilita el desarrollo y asegura la consistencia entre los diferentes proyectos y sus dependencias.
- **Manejo de Ficheros:** ITAINNOVA utiliza sistemas de almacenamiento avanzados para gestionar los ficheros de sus proyectos. Los ficheros pueden ser almacenados localmente en los servidores o en sistemas de almacenamiento en la nube para facilitar el acceso y la colaboración. Además, se emplean sistemas de control de versiones como GitLab para gestionar el historial de cambios en los ficheros y permitir la colaboración eficiente entre los miembros del equipo.

5.2.3. Tarea 2- Creación de ambientes de Conda

En esta fase, se llevará a cabo la configuración y creación de entornos virtuales utilizando la herramienta Conda. Estos ambientes virtuales asegurarán la consistencia y la correcta gestión de las dependencias del proyecto. Para una correcta utilización de la herramienta se utiliza el documento oficial de Conda Cheat Sheet y su guía de usuario [38][39].

Conda es una herramienta de gestión de paquetes y entornos que facilita la instalación y administración de librerías y dependencias necesarias para el desarrollo del proyecto. La creación de ambientes de Conda permite a los desarrolladores trabajar de manera aislada en un entorno limpio y coherente, lo que facilita la colaboración en equipo y el mantenimiento del proyecto a lo largo del tiempo. Además, la configuración de estos ambientes se puede replicar fácilmente en diferentes máquinas, lo que garantiza la reproducibilidad del desarrollo y los resultados obtenidos.

5.2.4. Tarea 3- Elaboración del Diagrama de clases para Bodyflow

En esta fase del proyecto, nos enfocamos en la creación de un Diagrama de Clases que refleje de manera visual la estructura y las relaciones existentes entre las clases que conforman la plataforma Bodyflow.

Este diagrama es una representación gráfica esencial para comprender cómo se organizan las clases dentro del sistema y cómo interactúan entre sí. A través de él, obtenemos una visión clara de las dependencias y responsabilidades de cada clase en el contexto general de la librería. El Diagrama de Clases desempeña un rol crucial en la comunicación interna del equipo, ya que facilita la discusión y comprensión colectiva sobre la arquitectura del proyecto. Además, este diagrama sirve como guía para futuras mejoras y ajustes en el código, ya que permite anticipar cómo los cambios pueden impactar en las interacciones entre las clases.

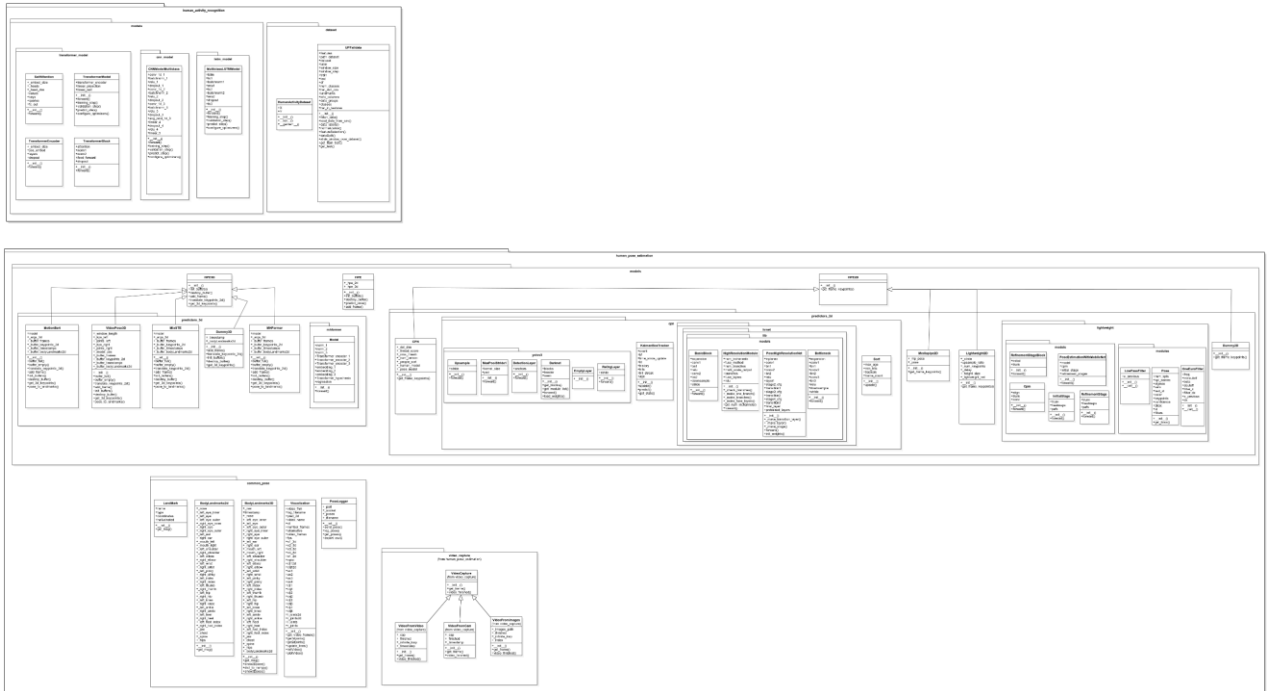


Figura 16: Diagrama de clases BodyFlow (versión pública-inicial)

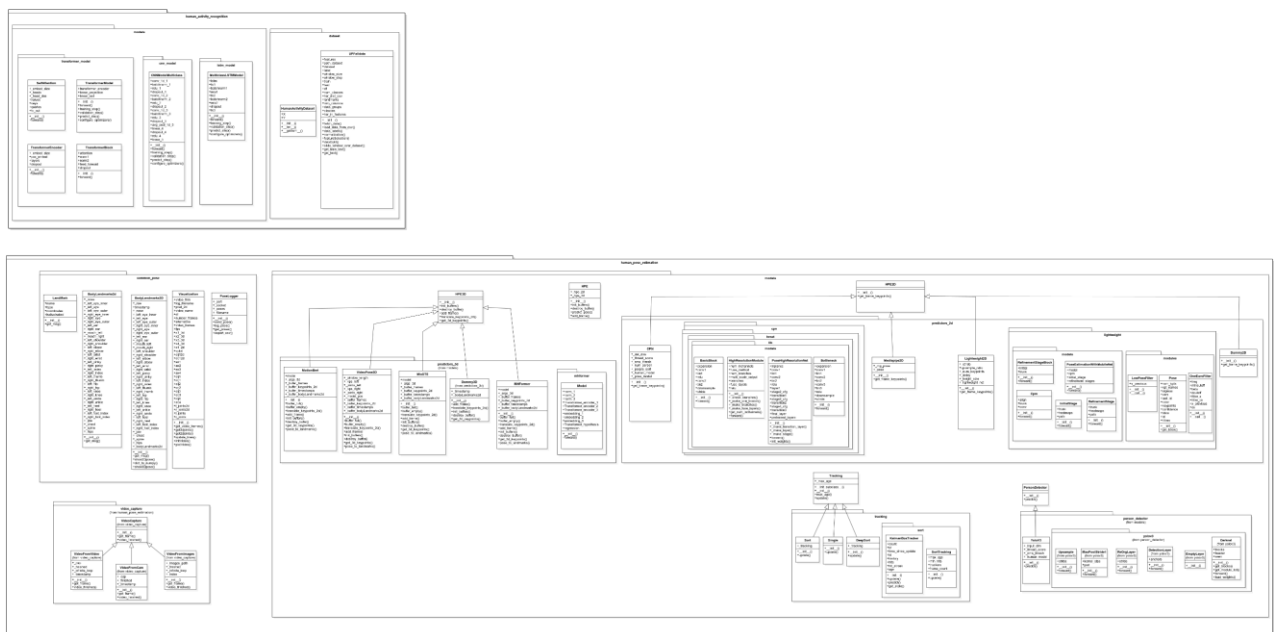


Figura 17: Diagrama de clases BodyFlow tras la aplicación del multitracking (versión en desarrollo)

El diagrama de clases y el conocimiento adquirido en el sprint anterior son recursos valiosos que permiten al desarrollador obtener una comprensión profunda de la arquitectura interna de la

librería. Los diagramas de las figuras 16 y 17 se adjuntan en alta resolución como material suplementario.

5.2.5. Tarea 4- Elaboración del Diagrama de clases para ExPose

Siguiendo con el mismo formato de la tarea anterior se desarrolla el diagrama de clases de ExPose, esta vez con la ayuda de la creación semiautomática de VisualParadigm. Dado que el código de ExPose es realmente complejo, la herramienta ha sido de gran utilidad al generar una versión inicial del diagrama que luego ha sido modificada y ajustada para obtener el mejor resultado posible.

VisualParadigm es una herramienta de modelado UML que facilita la creación de Diagramas de clases de manera visual y ágil. Al utilizar la creación semiautomática, la herramienta analiza el código fuente y genera un primer borrador del diagrama, mostrando las clases y sus relaciones detectadas automáticamente. Esta versión inicial ha servido como punto de partida para el desarrollo del diagrama, lo que ha ahorrado tiempo y esfuerzo en la etapa de diseño. El diagrama de la figura 18 se incluye en alta resolución en el material suplementario.

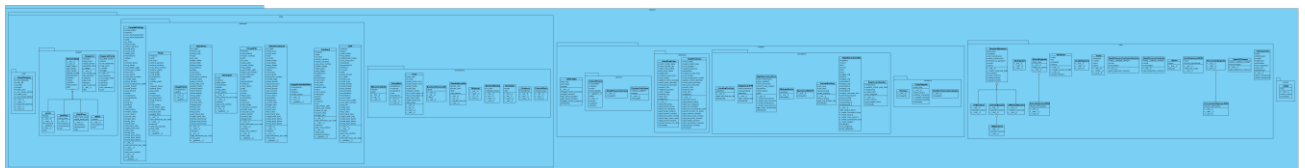


Figura 18: Diagrama de Clases ExPose

5.2.6. Tarea 5- Analizar ExPose para comprender su funcionamiento y exponerlo

El objetivo de esta tarea es realizar un análisis para obtener una visión clara y completa de la lógica interna de ExPose, incluyendo su estructura, algoritmos, flujos de datos y funcionalidades clave. Mediante la revisión del código fuente y la identificación de las diferentes partes y componentes, se comprende cómo la herramienta procesa la información, realiza cálculos y genera resultados.

Para lograr esto, ExPose sigue tres pasos fundamentales: procesamiento con R-CNN, SMPL-X y función de visualización.

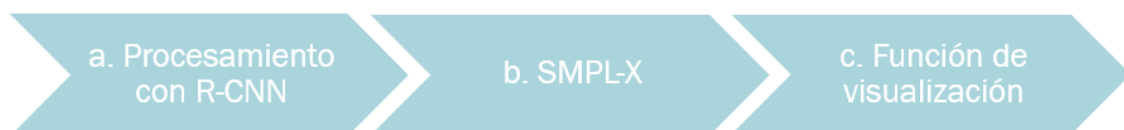


Figura 19: Esquema del funcionamiento de ExPose (Elaboración propia)

- a. RCNN (*Region-based Convolutional Neural Network*). En el primer paso, ExPose utiliza el enfoque RCNN, que incluye la detección de regiones de interés y la extracción de características visuales mediante una red neuronal convolucional. Fue propuesto por primera vez en el artículo "Rich feature hierarchies for accurate object detection and semantic segmentation" de Girshick et al. en 2014 [48]. Este enfoque consta de varias etapas:
- Región de interés (*ROI-Region Of Interest*). En esta etapa, se utilizan algoritmos de selección de regiones de interés, como Selective Search, para generar propuestas de regiones donde se sospecha que pueden estar presentes objetos. Estas regiones candidatas se extraen de la imagen original y se utilizan como entradas para la siguiente etapa.
 - Extracción de características. Cada región de interés propuesta se redimensiona para tener el mismo tamaño y se pasa por una red neuronal convolucional pre-entrenada. La red pre-entrenada se utiliza para extraer características visuales de las regiones de interés y convertirlas en vectores de características.
 - Clasificación de objetos. Los vectores de características de las regiones de interés se pasan a un clasificador SVM (*Support Vector Machine*) para determinar si cada región contiene un objeto de una clase específica o no. Por ejemplo, en nuestro caso, el clasificador SVM se entrena para distinguir entre regiones que contienen personas y regiones que no los contienen.
 - Regresión de cajas delimitadoras. Para mejorar la precisión de la detección, se utiliza un regresor para ajustar las cajas delimitadoras de las regiones propuestas y obtener una localización más precisa del objeto detectado.

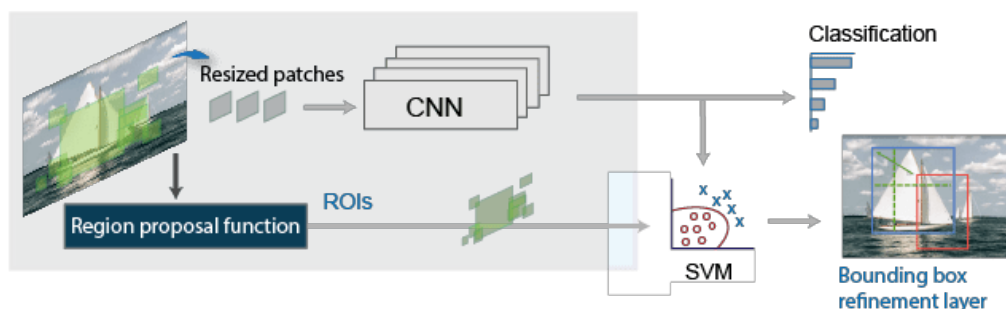


Figura 20: Esquema del funcionamiento de una R-CNN [30]

- b. SMPL-X. Una vez que se han extraído las características visuales con RCNN, ExPose utiliza el modelo SMPL-X para estimar la pose humana en 3D. Como ya se ha expuesto en el documento, este es un modelo avanzado que puede capturar una mayor variedad de

poses humanas y detalles anatómicos. Con la información de las características visuales, SMPL-X puede inferir la pose en 3D del sujeto en el vídeo.

- c. Función de visualización. Una vez que se ha estimado la pose humana en 3D, ExPose utiliza una función de visualización para mostrar los resultados de manera comprensible y amigable. Esta función genera representaciones visuales, en forma de modelos en 3D, que muestran la pose estimada del sujeto en el vídeo. Esto permite a los usuarios ver y analizar las poses humanas detectadas y obtener información valiosa sobre la dinámica del movimiento y la posición del cuerpo en el vídeo.

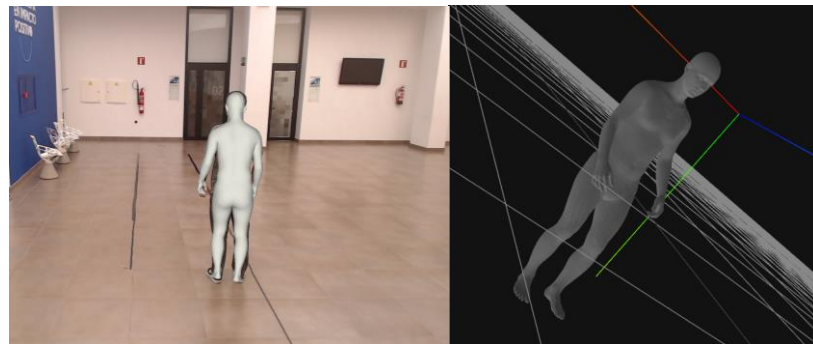


Figura 21: Resultados obtenidos con ExPose mediante la utilización de un vídeo realizado en ITAINNOVA

5.2.7. Tarea 6- Realización de una Demo de ExPose

Durante esta tarea, el enfoque se centra en desarrollar una demostración funcional y representativa de ExPose, que permita exhibir de manera práctica y visual las capacidades y características esenciales de la herramienta.



Figura 22: Flujo en el tratamiento de las imágenes que se quieren procesar y la correspondiente incorporación a la demo

El formato de salida de BodyFlow se presenta en un vídeo, y la propuesta es incorporar la salida de ExPose en este mismo vídeo. Al hacerlo, se logra una visualización conjunta y más completa de los resultados de ambas herramientas en una presentación unificada.

Esta integración servirá como un poderoso reclamo del trabajo de integración que se tiene previsto realizar, permitiendo a los usuarios apreciar de manera directa y convincente la

capacidad combinada de ambas herramientas en la estimación de poses humanas y el análisis de movimientos.

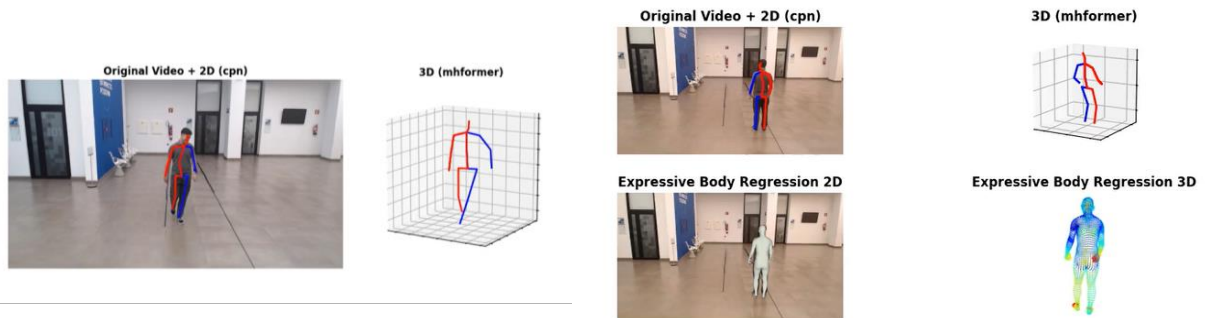


Figura 23: Imagen del vídeo obtenido con Bodyflow mediante la utilización de MHFormer

Figura 24: Imagen del vídeo demo obtenido utilizando la salida de Bodyflow+ MHFormer y ExPose por separado

5.2.8. Evaluación del sprint

El segundo sprint ha sido completado con éxito, sin embargo, la tarea 5 "Análisis ExPose para comprender su funcionamiento y exponerlo", se ha realizado a más alto nivel del esperado por el Scrum Master. Como resultado, esta tarea se reincorporará al tablero Kanban para ser abordada en el siguiente sprint. Se considera por tanto que se ha realizado un 92% del trabajo esperado en 65 horas de esfuerzo sobre las 60 que habían sido previstas.

5.3. *Sprint 3: Integración de ExPose y visualización de resultados*

5.3.1. Configuración tablero Kanban

En este sprint se va a trabajar sobre los objetivos *Integración de un nuevo módulo "ExPose" en la librería de Bodyflow y Visualización propia*. El tiempo estimado es de 90 horas.

<u>Tarea 0</u>	Analizar ExPose para comprender su funcionamiento y exponerlo
<u>Tarea 1</u>	Estudio de las licencias de ExPose
<u>Tarea 2</u>	Redacción y exposición de la planificación de integración
<u>Tarea 3</u>	Implementación
<u>Tarea 4</u>	Creación de un <i>wrapper</i> (envoltorio) que facilite la instalación
<u>Tarea 5</u>	Corrección de <i>paths</i> propio de este tipo de integración
<u>Tarea 6</u>	Comprobación de funcionamiento a través de tests unitarios
<u>Tarea 7</u>	<i>Profiling</i> (perfilado) del código
<u>Tarea 8</u>	Gráfico del flujo de funcionamiento de ExPose tras su integración
<u>Tarea 9</u>	Diseño de los formatos de salida de los datos generados
<u>Tarea 10</u>	Visualización de la salida a través de un vídeo integrado

5.3.2. Tarea 0- Analizar ExPose para comprender su funcionamiento y exponerlo

Tras recibir el feedback del Scrum Master, esta vez nos enfocamos especialmente en el funcionamiento de SMPL-X dentro de ExPose. En la primera exposición, SMPL-X se había tratado como una caja negra, pero ahora se requiere una comprensión más profunda de su funcionamiento interno.

Se propone desentrañar los detalles y características de SMPL-X para comprender cómo contribuye al proceso general de estimación de poses humanas en 2D y 3D dentro de ExPose. Para lograrlo, llevamos a cabo varias iteraciones de análisis detallado paso a paso. Mediante estas barridas "*step by step*", desglosamos y examinamos minuciosamente cada aspecto de SMPL-X y su interacción con el resto de componentes.



Figura 25: Flujo de funcionamiento de SMPL-X (Elaboración propia)

De esta forma, SMPL-X se divide en:

- a. Backbone (Red Neuronal Convolutacional). En primer lugar, se utiliza una *backbone*, que es una red neuronal convolutacional, para procesar las imágenes de entrada y extraer características relevantes del cuerpo humano. Esta *backbone* es responsable de identificar patrones, formas y estructuras del cuerpo en las imágenes.

Dentro del marco de ExPose, la elección de la *backbone* se centra en la implementación de HRNet (*High-Resolution Network*), una arquitectura de red neuronal convolutacional que ha demostrado ser altamente efectiva en aplicaciones de visión por computadora que requieren una alta resolución espacial y una detección precisa de detalles minuciosos.

A diferencia de las arquitecturas convencionales de redes convolucionales, que tienden a disminuir progresivamente la resolución de las características a medida que avanzan en las capas, HRNet adopta una aproximación distinta. Mantiene múltiples niveles de resolución de características en cada etapa de la red, lo que da como resultado una capacidad única para conservar detalles cruciales y capturar características de alto nivel junto con detalles finos. Esta capacidad de HRNet para retener información de alta resolución y a la vez extraer detalles de nivel micro ofrece una representación visual más completa y precisa de los datos.

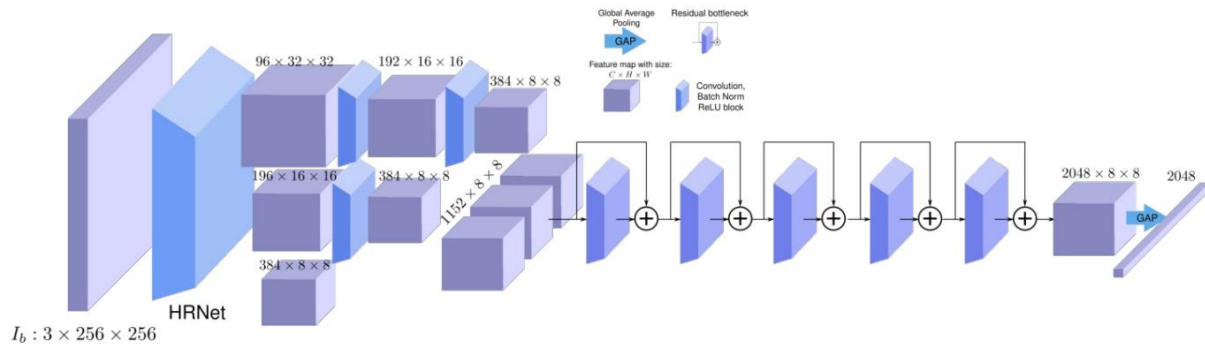


Figura 26: Estructura del extractor de características utilizado por la red de predicción del cuerpo. La imagen I_b se introduce en HRNet para extraer mapas de características de múltiples escalas. Luego, estos mapas se procesan mediante bloques convolucionales adicionales y se reducen de tamaño a la misma resolución espacial. Todos los mapas de características se concatenan y se ingresan a 5 bloques residuales seguidos de una operación de promedio global que produce el vector de características final F_b [42].

- b. *Regressor*. Una vez que las características del cuerpo han sido obtenidas por la *backbone*, se utilizan como entrada para el *regressor*. El *regressor* es otro componente de la red neuronal que realiza un proceso de regresión para obtener los parámetros específicos del cuerpo. Estos parámetros son utilizados para estimar la pose del cuerpo en 2D y 3D. Se configura y crea un regresor MLP, (Figura 27) con una dimensión de entrada basada en las características extraídas por la columna vertebral de la red, y una dimensión de salida basada en el tamaño de los parámetros esperados.

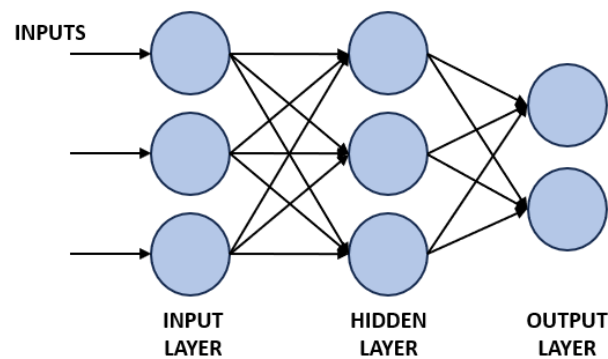


Figura 27: MLP hace referencia a "Multi-Layer Perceptron" (Perceptrón Multicapa), que es un tipo de red neuronal artificial que consta de múltiples capas de nodos (neuronas). Cada capa está conectada a la siguiente mediante conexiones ponderadas. Los nodos en las capas ocultas utilizan una función de activación no lineal para procesar la información y pasarla a la siguiente capa. (Elaboración propia)



- c. *Body Model* (Modelo del Cuerpo). Por último, se utiliza el modelo del cuerpo para representar y visualizar la pose estimada en 2D y 3D. Este modelo incluye información sobre la posición de las articulaciones, la orientación de las extremidades y otros detalles anatómicos del cuerpo. Es notable la utilización de diversas técnicas, incluyendo "*Linear Blend Skinning*" (LBS), que se emplea con el propósito de mejorar la simulación del movimiento de las articulaciones.

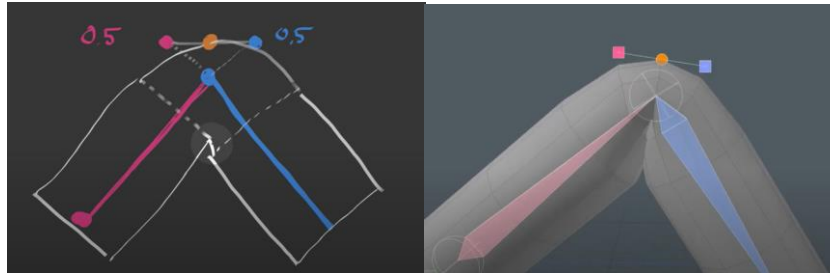


Figura 28: Estructuras esqueléticas utilizadas para deformar la malla con el fin de obtener resultados más realistas en el movimiento de las articulaciones. [43]

5.3.3. Tarea 1 - Estudio de las licencias de ExPose

En esta tarea, se realiza un análisis detallado de las licencias asociadas a ExPose, con el objetivo de comprender los términos y restricciones legales que rigen su uso, distribución y modificaciones. Esto es esencial para asegurarse de que el proyecto cumple con las condiciones establecidas por las licencias y evitar posibles problemas legales en el futuro. La licencia de ExPose establece una serie de términos y condiciones que regulan el uso y manejo de los datos, modelos y software proporcionados. Aquí se expone un resumen detallado de los aspectos clave de esta licencia:

- Propósito limitado. Los datos y software de ExPose pueden ser utilizados únicamente para llevar a cabo investigaciones científicas no comerciales, actividades educativas no comerciales o proyectos artísticos no comerciales.
- Distribución restringida. La distribución de los datos y software está limitada. No se permite la copia, compartición, distribución, reventa, sublicencia o transferencia de estos, a menos que sea con fines de archivo. Cualquier distribución adicional debe contar con el permiso previo por escrito de Max-Planck.
- Uso responsable. Los usuarios deben utilizar los datos y software de manera responsable y ética. No pueden ser utilizados para crear contenido falso, difamatorio, engañoso o degradante. Además, no se puede utilizar el software para entrenar métodos, algoritmos o redes neuronales con fines comerciales, militares, pornográficos o difamatorios.

- d. Garantías limitadas. Los datos y software se proporcionan sin ninguna garantía expresa o implícita. El licenciante no se hace responsable de posibles daños causados por el uso de estos. Los usuarios deben entender que pueden contener errores y que su uso está bajo su propio riesgo.
- e. Limitación de responsabilidad. El licenciante solo es responsable por daños causados intencionalmente o por negligencia grave, según lo permitido por la ley. No hay obligación de proporcionar servicios de mantenimiento, actualizaciones ni correcciones de errores.
- f. Referencia obligatoria. Si los datos y software se utilizan en publicaciones o trabajos, los usuarios deben hacer referencia a un artículo específico proporcionado en la licencia.
- g. Gobernanza legal. El acuerdo de licencia está sujeto a las leyes de la República Federal de Alemania, a excepción de la Convención de Ventas de las Naciones Unidas.

En resumen, la licencia de ExPose establece claramente el alcance y las limitaciones del uso de los datos, modelos y software. Está diseñada para fomentar un uso responsable y ético en el ámbito de la investigación científica, educación y proyectos artísticos no comerciales, mientras que prohíbe explícitamente usos inapropiados o comerciales sin autorización.

5.3.4. Tarea 2- Redacción y exposición de la planificación de integración

En esta tarea, se procede a elaborar y presentar la planificación detallada para la integración del modelo ExPose en el entorno de desarrollo de Bodyflow. Tal como se ha observado en los estudios anteriores, BodyFlow opera en un formato modular, lo que implica que es necesario realizar una adaptación del modelo para que se integre de manera coherente con las funciones que se utilizan en la librería.

Primero se establecen objetivos claros para la integración. Estos son mejorar las métricas de la estimación de la pose 3D en el proyecto utilizando los conceptos de SMPL y SMPL-X, e incorporar la capacidad de captura volumétrica del cuerpo, la cara y las manos.

Se realiza el diseño de la integración, se elabora un diseño exhaustivo sobre la manera en que ExPose se incorporará en el proceso de trabajo de Bodyflow. Esto conlleva la generación de un módulo adicional específico "*expose*"; así como la creación de un nuevo archivo de Python derivado de HPE3D "*Expose.py*". Además, se ajustan las funciones ya existentes para que puedan incluir los parámetros necesarios para el uso de ExPose, incluyendo la imagen correspondiente.

Finalmente, se elabora una exposición que sintetiza la planificación de la integración y se comparte con el Scrum Master y Carlos Maraños, quien desempeña el rol de técnico I+D+I en Big Data en ITAINNOVA. Esto garantiza una comprensión mutua y un compromiso conjunto con la puesta en práctica del plan facilitando la integración y dando soporte al mismo.

5.3.5. Tarea 3- Implementación

El objetivo principal es llevar a cabo la implementación de la integración de ExPose. Tal como se explicó en la actividad previa, se avanza con la inclusión de un nuevo módulo denominado "expose", el cual se deriva del código original. Además, como es lógico se desarrolla un código específico desde cero, basado en la estructura de HPE3D, con el propósito de permitir el uso eficaz de ExPose. También se efectúan ajustes menores en el "inference_server", el main del módulo HPE, con la finalidad de que los usuarios puedan aprovecharlo de manera similar a cómo interactúan con los modelos previamente establecidos en el sistema.

5.3.6. Tarea 4- Creación de un wrapper (envoltorio) que facilite la instalación

En esta tarea, el trabajo se enfoca en crear un envoltorio (*wrapper*) con el propósito de simplificar el proceso de instalación de ExPose dentro del contexto de Bodyflow. Este envoltorio, presentado en forma de archivo "readme", está diseñado para guiar al usuario en la puesta en marcha del nuevo módulo. Proporciona instrucciones claras sobre cómo descargar e integrar ExPose de manera efectiva.

Al proporcionar un envoltorio de este tipo, se pretende facilitar el uso adecuado del módulo en conformidad con la licencia que se ha presentado anteriormente. Esto permite que cada usuario utilice ExPose de manera responsable y legal, sin que sea necesario realizar ninguna distribución adicional por nuestra parte. El envoltorio brindará una descripción detallada de los pasos a seguir, desde la descarga del módulo hasta su inserción en el entorno de Bodyflow.

5.3.7. Tarea 5- Corrección de paths propio de este tipo de integración

El código obtenido de ExPose no es sujeto de modificaciones, ya que, como se ha explicado previamente, los usuarios lo descargarán en su forma original. Por lo tanto, nuestra tarea es ajustar nuestro propio código para que se integre como un módulo sin enfrentar problemas relacionados con la duplicación de nombres de rutas. En este punto, además de corregir ciertas rutas existentes para convertirlas en rutas completas, es necesario efectuar modificaciones en el archivo de configuración "conf.yaml". Esto se debe a que en Bodyflow ya existe una carpeta denominada "data" utilizada para guardar los inputs y outputs de ejecución, y una de las carpetas que necesitamos incorporar posee el mismo nombre. Por lo tanto, se toma la decisión de agregar esta carpeta a la estructura de la carpeta "modelos". Esta elección se basa en que la carpeta contiene los modelos necesarios para ExPose, y esta acción evita la duplicación de nombres de carpetas.

Los cambios realizados en el archivo de configuración reflejan esta modificación. Estos ajustes permiten asegurar una organización coherente y evitar posibles conflictos entre nombres de carpetas, garantizando que la integración de ExPose en Bodyflow sea eficiente y libre de ambigüedades.

5.3.8. Tarea 6- Comprobación de funcionamiento a través de tests unitarios

En esta etapa, se lleva a cabo la tarea de verificar el funcionamiento del sistema mediante la ejecución de pruebas unitarias. Estas pruebas tienen como objetivo evaluar cada componente individual de ExPose que ha sido integrado en el entorno de Bodyflow, asegurando que su funcionamiento sea coherente y libre de errores.

Se diseñan estos casos de prueba para abarcar diferentes situaciones y escenarios posibles, con el fin de poner a prueba la robustez y confiabilidad de la integración. La ejecución de pruebas unitarias utilizando la biblioteca "unittest" nos permite identificar cualquier problema o fallo en el funcionamiento de ExPose dentro del contexto de Bodyflow. Sometemos cada componente a una serie de entradas y condiciones diversas para evaluar su respuesta y comportamiento. Si detectamos problemas, realizamos los ajustes necesarios para garantizar un funcionamiento adecuado.

Además, en esta tarea se lleva a cabo la actualización de los requisitos (*requirements*) y sus respectivas licencias. Se procede a la creación de un entorno completamente nuevo y se verifica su funcionalidad, asegurando así que todas las librerías estén documentadas y sean compatibles con el proyecto.

5.3.9. Tarea 7- Profiling (perfilado) del código

En esta fase, estamos llevando a cabo la tarea de perfilado utilizando la herramienta "SnakeViz". El objetivo principal es analizar el rendimiento y la eficiencia del sistema integrado a través de la visualización detallada de los perfiles de ejecución generados por SnakeViz.

Durante esta tarea, ejecutamos el sistema integrado mientras SnakeViz recopila datos sobre el tiempo de ejecución y el uso de recursos de cada componente. Esta información se presenta de manera gráfica y visual, lo que nos permite identificar fácilmente los puntos críticos de rendimiento y las áreas que requieren optimización. Se explora de manera interactiva los perfiles de ejecución y se profundiza en los detalles de las funciones y los procesos que consumen más tiempo y recursos. Esto nos ayuda a entender cómo se distribuye el tiempo de ejecución y a identificar posibles ineficiencias. Se lleva a cabo tanto en vídeos de duración breve como extensa, con el propósito de abordar diversas situaciones. Es importante destacar que, en el caso de vídeos extensos, la mayor parte del proceso de ejecución se concentra en el procesamiento de imágenes, lo que dificulta la identificación de posibles puntos de congestión o ralentización.

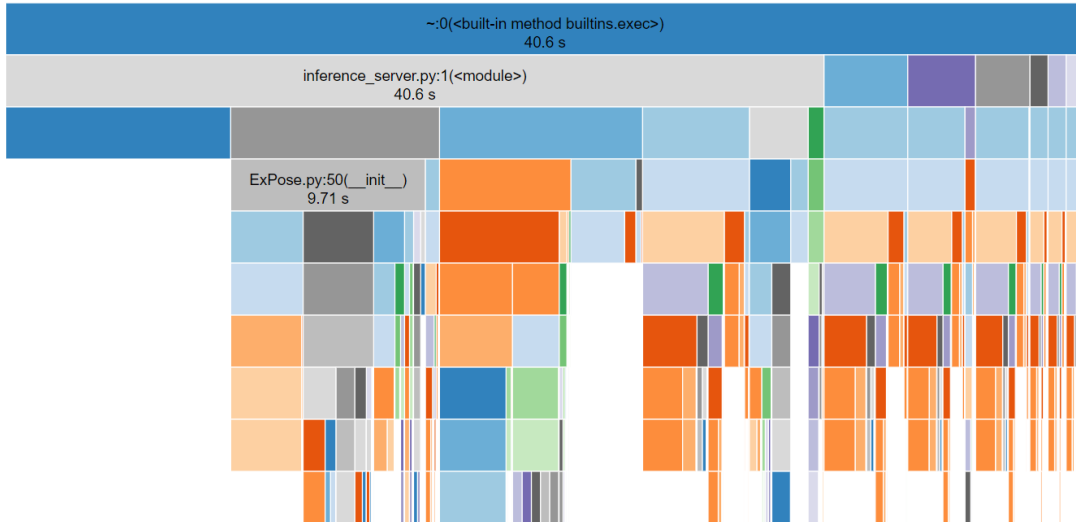


Figura 29: Captura de pantalla de Snakeviz que muestra los tiempos de ejecución detallados en la ejecución del programa

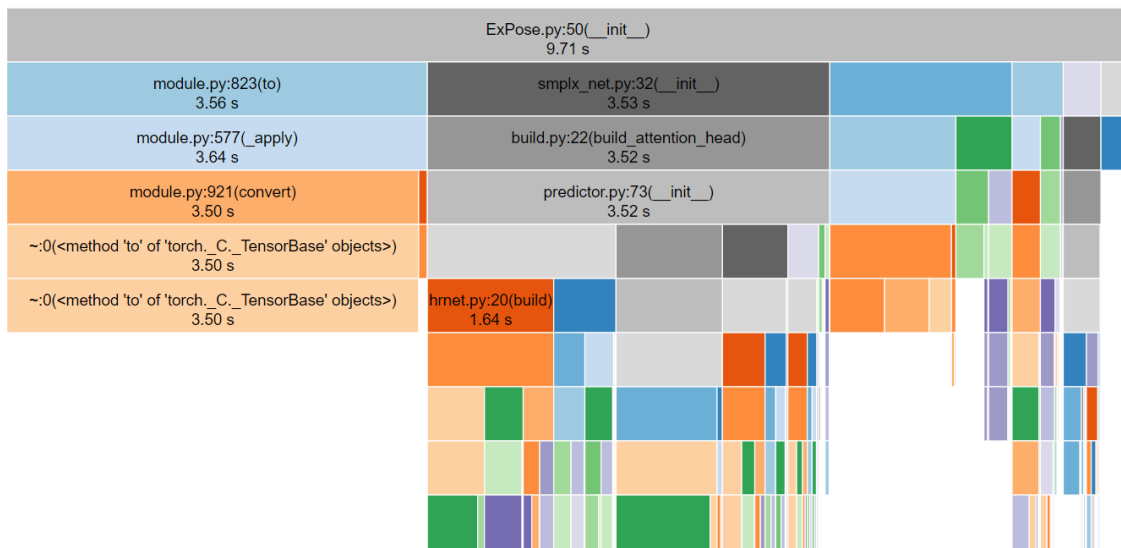


Figura 30: Captura de pantalla de Snakeviz con el desglose de los tiempos de ejecución durante el análisis de rendimiento de "Expose.py"

5.3.10. Tarea 8- Gráfico del flujo de funcionamiento de ExPose tras su integración

En esta tarea se aborda la creación de un gráfico que ilustre el flujo de funcionamiento de ExPose después de su integración en el entorno de Bodyflow. El objetivo es proporcionar una representación visual clara y comprensible de cómo interactúan los diferentes componentes y del modelo en el marco de la librería.

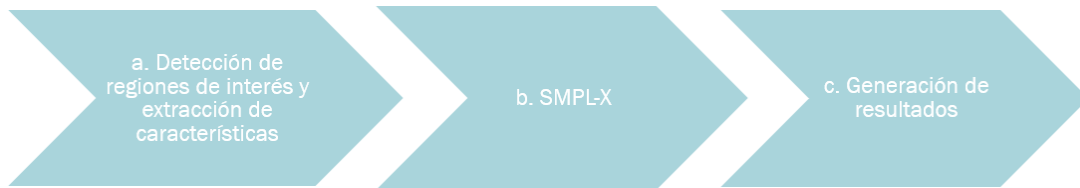


Figura 31: Flujo-resumen del funcionamiento de ExPose tras su integración (Elaboración propia)

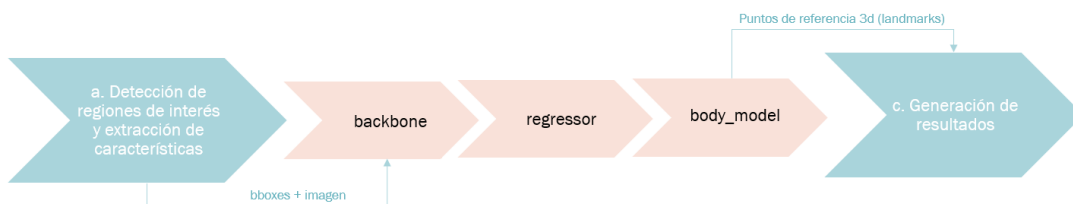


Figura 32: Flujo del funcionamiento de ExPose tras su integración (Elaboración propia)

- En esta fase, se lleva a cabo la detección de áreas de interés en las imágenes y la extracción de características relevantes. En vez de utilizar la arquitectura RCNN que se explicó en la tarea 5 de este sprint, se opta por el enfoque modular de la librería en el que se incorpora YOLOv3 de forma predeterminada. YOLOv3, un modelo de detección de objetos que divide la imagen en cuadrículas y asigna una caja delimitadora y una clase a cada cuadrícula.
- Una vez obtenidos los resultados de detección, se procede a la utilización de la versión ExPose de SMPL-X. En esta fase, se emplean las cajas delimitadoras (*bboxes*) obtenidas previamente, así como la imagen en sí misma. Estos elementos actúan como insumos para el proceso de estimación que se lleva a cabo en esta fase.
- Finalmente, se genera un conjunto de resultados enfocándose exclusivamente en los datos de interés. Este proceso implica la extracción y presentación de los resultados que son relevantes para el objetivo de la integración. Se asignan los puntos de referencia relevantes, y el resto se ignoran, consiguiendo que funcione en la librería como el resto de los modelos en su representación 3d. Adicionalmente, se registra la malla 3d correspondiente para su posterior utilización.

5.3.11. Tarea 9- Diseño de los formatos de salida de los datos generados

Esta tarea implica definir la disposición y organización de la información relevante, así como determinar los tipos de archivos o representaciones que se utilizarán para la entrega de los resultados obtenidos. Con el propósito de mantener una coherencia con la estructura de la biblioteca Bodyflow, se opta por conservar la generación de un vídeo de salida y añadir otro que

muestre las representaciones generadas por ExPose. Se valora la importancia de mantener una presentación de información ordenada y coherente por lo que tras varias opciones se decide seguir la disposición de la *Figura 33*.

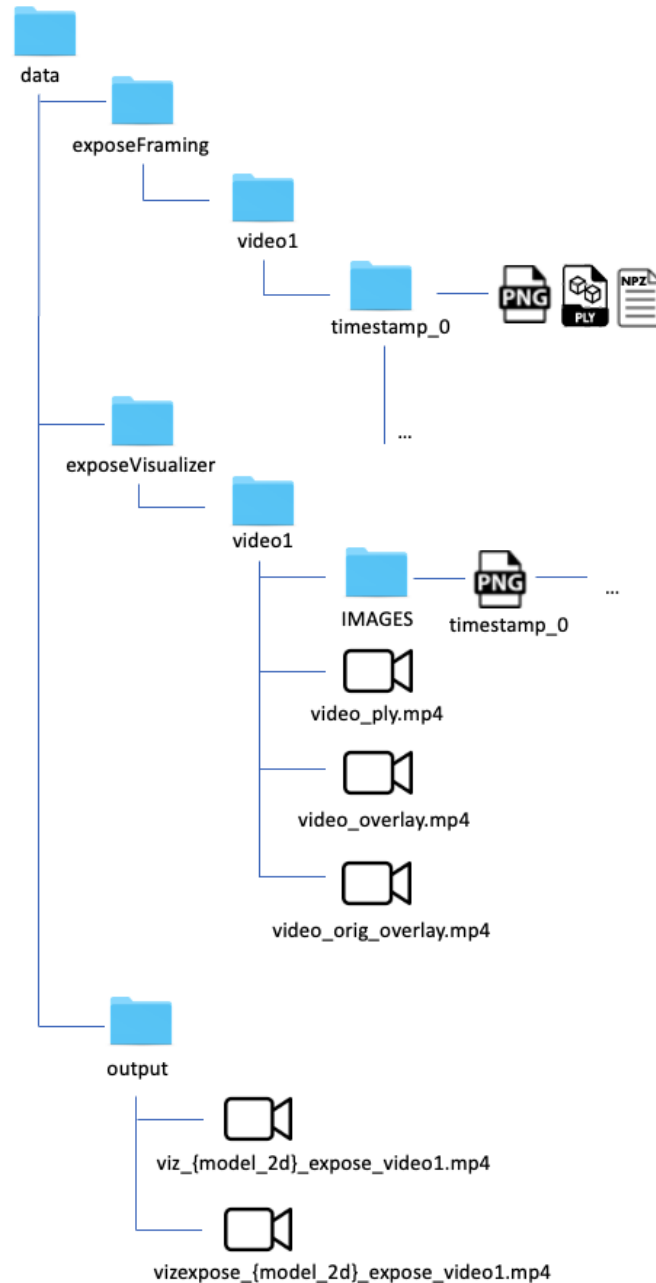


Figura 33: La salida generada de forma predeterminada en cada frame se almacena en el directorio "exposeFraming". Aquí, se encuentran disponibles tanto las imágenes originales como sus versiones modificadas, junto con un archivo de información en formato .npz y las representaciones volumétricas en archivos .ply. Paralelamente, en la carpeta "exposeVisualizer" se archivan las imágenes derivadas de la malla, así como los vídeos generados a partir de estas imágenes y las obtenidas de manera predeterminada. Finalmente, en la carpeta "output", se conserva el vídeo resultante generado por defecto en Bodyflow, y se agrega otro que fusiona los dos vídeos más representativos de la salida generada por ExPose. (Elaboración propia)

5.3.12. Tarea 10- Visualización de la salida a través de un vídeo integrado

En esta tarea, se aborda la creación de un vídeo integrado que muestre la salida generada por ExPose de manera visual, el denominado `vizexpose_{model_2d}_expose_{videoname}.mp4` en la tarea anterior. El objetivo principal es presentar de manera efectiva los resultados obtenidos a través de un formato audiovisual que permita una comprensión más clara y concisa de la funcionalidad y utilidad de ExPose en el contexto de Bodyflow.

Siguiendo el formato del vídeo que se estaba generando y el feedback de la demo que se creó en el sprint anterior se adapta la salida. Se puede ver el resultado en la *Figura 34* y *Figura 35*.

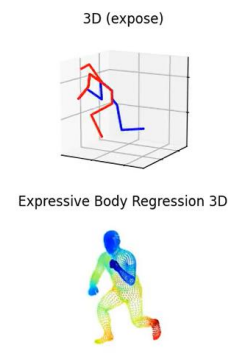
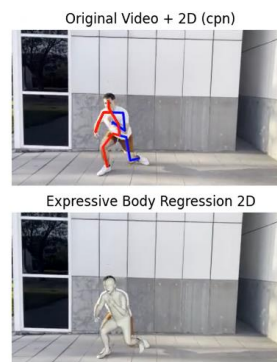
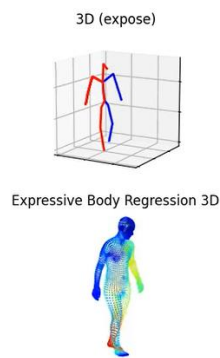
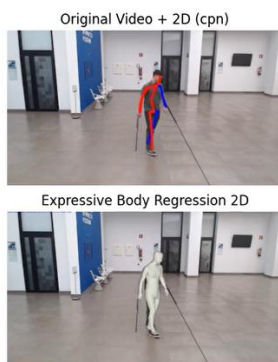


Figura 34: Imagen que muestra el vídeo integrado resultante de la ejecución de ExPose en un vídeo de prueba realizado en el interior.

Figura 35: Imagen que muestra el vídeo integrado resultante de la ejecución de ExPose en un vídeo de prueba realizado en el exterior.

5.3.13. Evaluación del sprint

La evaluación de este sprint se inicia con una presentación adicional sobre el análisis de ExPose, que concuerda con las expectativas del Scrum Master. Luego de examinar las licencias y diseñar la planificación de integración, se recibe un feedback positivo. Al finalizar, tras llevar a cabo la implementación y abordar las demás tareas, se verifica que han sido realizadas exitosamente y se concluye el sprint. Es notable mencionar que se realizó un esfuerzo de 86 horas sobre las 90 horas previstas.

5.4. *Sprint 4: Módulo de sincronización*

5.4.1. *Configuración tablero Kanban*

En este sprint se va a trabajar sobre el objetivo *Actualización de la función de sincronización*. El tiempo estimado es de 60 horas.

<u>Tarea 1</u>	Análisis del funcionamiento módulo de sincronización
<u>Tarea 2</u>	Adaptar el funcionamiento a la nueva funcionalidad <i>multipose</i>
<u>Tarea 3</u>	Adaptar al cambio en la denominación de los tiempos (<i>timestamps</i>)
<u>Tarea 4</u>	Adaptar ExPose para su funcionamiento en serie
<u>Tarea 5</u>	Ejecución y sincronización del "FALL-UP dataset" con ExPose

5.4.2. *Tarea 1- Análisis del funcionamiento módulo de sincronización*

En esta tarea, se lleva a cabo un análisis detallado y exhaustivo del funcionamiento del módulo de sincronización, profundizando en sus características y comportamiento. La biblioteca BodyFlow implementa una función de sincronización que se encarga de realizar una fusión de datos a nivel de entrada. En primer lugar, como se ha expuesto el Módulo 1 (HPE) extrae la pose de los fotogramas capturados por múltiples cámaras de vídeo, y solo entonces esta pose se combina con la información en bruto de los sensores inerciales. Una vez que la información está correctamente fusionada, se pasa al clasificador en el Módulo 2 (HAR).

Los tres pasos de la función de sincronización se describen de la siguiente manera:

- Se sincronizan todos los sensores entre sí.
- Los vídeos se sincronizan utilizando la marca de tiempo.
- Ambos conjuntos de datos se emparejan y se guardan en un único archivo .csv.

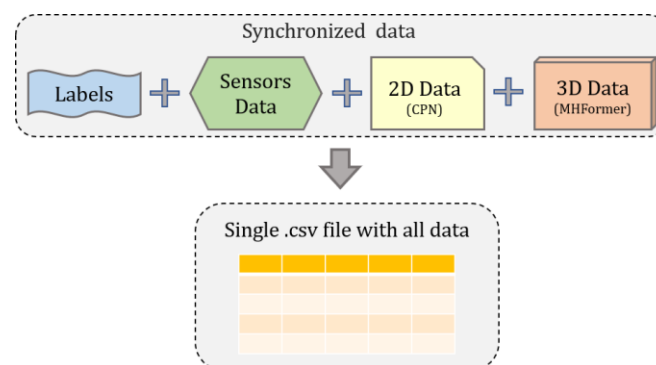


Figura 36: Estructura de la función de sincronización [44]

5.4.3. Tarea 2- Adaptar el funcionamiento a la nueva funcionalidad multipose

En esta tarea, se realiza la modificación del funcionamiento existente para incorporar la nueva funcionalidad de *multi*pose en el módulo de sincronización. Esta nueva característica permite manejar y fusionar datos de múltiples poses simultáneamente, lo que amplía la capacidad del sistema para capturar y procesar información de varios usuarios al mismo tiempo.

5.4.4. Tarea 3- Adaptar al cambio en la denominación de los tiempos (timestamps)

En esta tarea, se actualiza funcionamiento existente para acomodar el cambio en la denominación de los tiempos (*timestamps*) en el módulo de sincronización. Este reajuste es necesario debido a actualizaciones en la forma en que se registran los tiempos en los nombres de archivo que se obtienen del Módulo 1. La tarea implica modificar el código correspondiente a la función de sincronización para asegurarse de que los nuevos tiempos se gestionen correctamente y se integren en el proceso de fusión de datos de manera precisa y coherente.

5.4.5. Tarea 4- Adaptar ExPose para su funcionamiento en serie

En esta tarea, se aborda la adaptación de ExPose para que pueda funcionar en modo secuencial o en serie, lo que implica permitir el procesamiento de múltiples secuencias de vídeo de manera continua y eficiente. Hasta este punto, ExPose se ha centrado en la generación de resultados para un solo vídeo a la vez. Sin embargo, en muchas aplicaciones prácticas, es necesario procesar varios vídeos uno tras otro sin interrupciones como es el caso de este módulo de sincronización.

5.4.6. Tarea 5- Ejecución y sincronización del "FALL-UP dataset" con ExPose

En esta tarea, se aborda la ejecución y sincronización del conjunto de datos "FALL-UP" con ExPose y la función de sincronización actualizada. Este conjunto de datos está disponible públicamente y está creado y mantenido por la Universidad Panamericana [29]. Este conjunto de datos multimodal incluye 11 actividades diferentes realizadas por 17 individuos jóvenes, donde cada actividad es repetida tres veces por cada sujeto.

El "FALL-UP dataset" es utilizado como entrada con el objetivo de analizar y obtener resultados relevantes sobre la detección de posturas y movimientos en las secuencias de vídeo presentes en el conjunto de datos. Durante la ejecución de la tarea se ha enfrentado un desafío relacionado con el coste computacional. Se ha observado que este proceso ha requerido una considerable cantidad de recursos de cómputo y tiempo para completarse.

En concreto, el tiempo dedicado a esta tarea ha abarcado aproximadamente una semana y media (253.2 horas). Durante este período, los recursos de procesamiento, memoria y tiempo de ejecución se han empleado de manera intensiva para realizar la estimación de pose en las secuencias de vídeo, así como para el emparejamiento de la información en un único csv.

5.4.7. Evaluación del sprint

En este cuarto sprint se realiza un esfuerzo de 56 horas sobre las 60 previstas completando con éxito todas las tareas planteadas, las cuales fueron validadas por la supervisión del Scrum Master.

No se considera en este cálculo el tiempo computacional que, aunque fue un factor importante en la ejecución del proyecto, no ha impedido su realización.

5.5. *Sprint 5: Evaluación del modelo y resultados*

5.5.1. Configuración tablero Kanban

En este sprint se va a trabajar sobre los objetivos *Evaluación situacional de los algoritmos y Utilización del módulo de HAR para la obtención de métricas*. El tiempo estimado es de 60 horas.

<u>Tarea 1</u>	Análisis del funcionamiento módulo de HAR
<u>Tarea 2</u>	Ejecutar las diferentes combinaciones de modelos HAR y características que sean útiles para la evaluación de ExPose
<u>Tarea 3</u>	Presentación y análisis de resultados

5.5.2. Tarea 1- Análisis del funcionamiento módulo de HAR

El análisis del funcionamiento del módulo de Reconocimiento de Actividades Humanas (HAR) se refiere a la evaluación detallada de cómo este componente dentro del proyecto opera y desempeña su tarea específica. Se involucra en la tarea la comprensión de cómo se recogen los datos de entrada, cómo se procesan y cómo se utilizan para predecir y clasificar las actividades humanas. Además, se evalúa la precisión y eficacia de las técnicas de aprendizaje automático utilizadas en términos de su capacidad para reconocer con precisión las actividades.

En esta tarea, también se cuenta con el respaldo de Irene López Técnica I+D+I en Big Data, quien está disponible para ayudar a resolver cualquier duda relacionada con la toma de decisiones sobre la arquitectura del sistema y en especial para proporcionar información actualizada sobre el estado y funcionamiento del módulo en cuestión. Su experiencia y conocimiento contribuyen a asegurar que las decisiones tomadas sean fundamentadas y que el módulo funcione de manera óptima en el contexto del proyecto.

5.5.3. Tarea 2- Ejecutar las diferentes combinaciones de modelos HAR y características que sean útiles para la evaluación de ExPose

La biblioteca BodyFlow fue evaluada con el conjunto de datos FALL-UP como se ha podido observar en su utilización para la sincronización. En particular, la biblioteca fue evaluada utilizando los datos de este conjunto de datos provenientes de cinco sensores inerciales y dos cámaras. Además, para esta evaluación, se utilizaron los ensayos 1 y 2 para el entrenamiento, mientras que el ensayo 3 se utilizó para las pruebas.

El *Anexo III* muestra los resultados de las pruebas de los experimentos en el conjunto de datos para diferentes combinaciones de modelos HAR y tipos de características. La biblioteca permite el uso de diferentes combinaciones de características para estudiar y analizar qué tipo de sensor es más adecuado para reconocer qué tipo de actividad. Específicamente, la columna "Características" de la Tabla 2 muestra qué entradas se han utilizado para el modelo HAR: a) 3D:

pose en 3D del módulo 1, b) 2D: pose en 2D del módulo 1, c) IMUs: los cinco sensores inerciales, d) "ankle": solo el sensor inercial colocado en el tobillo y e) "all": una fusión de pose en 2D, pose en 3D y los cinco sensores inerciales.

En particular, estos resultados se obtuvieron utilizando el algoritmo de pose CPN en 2D, el algoritmo de pose MHFormer en 3D y un tamaño de ventana de 41 fotogramas.

Para la realización de la comparativa de las métricas con ExPose se ha procedido con la ejecución utilizando los mismos parámetros, el algoritmo de pose CPN en 2D y un tamaño de ventana de 41 fotogramas. Se ha ejecutado tanto con la entrada 3d de ExPose como con "all" para poder comparar los resultados de manera global. De esta forma se incorporan los resultados englobándose en la *Tabla 1*

HAR Model	Features	F1	Accuracy	Precision
LSTM	all (MHFormer)	0.718	0.848	0.738
LSTM	all (ExPose)	0.781	0.937	0.797
LSTM	3d (MHFormer)	0.787	0.888	0.795
LSTM	3d (ExPose)	0.746	0.893	0.767
LSTM	2d	0.692	0.864	0.702
LSTM	imus	0.702	0.916	0.718
LSTM	ankle	0.515	0.768	0.550
CNN	all (MHFormer)	0.781	0.892	0.816
CNN	all (ExPose)	0.761	0.937	0.763
CNN	3d (MHFormer)	0.760	0.836	0.797
CNN	3d (ExPose)	0.817	0.913	0.833
CNN	2d	0.616	0.835	0.676
CNN	imus	0.655	0.833	0.685
CNN	ankle	0.664	0.795	0.699
Transformer	all (MHFormer)	0.865	0.891	0.882
Transformer	all (ExPose)	0.803	0.944	0.836
Transformer	3d (MHFormer)	0.825	0.881	0.818
Transformer	3d (ExPose)	0.836	0.931	0.827
Transformer	2d	0.825	0.864	0.819
Transformer	imus	0.793	0.896	0.852
Transformer	ankle	0.651	0.785	0.679
Espinosa et al. [45]		0.7294	0.8226	0.7425
Martinez-Villasenor et al.[29]		0.712	0.951	0.718
Suarez et al. [46]		0.8367	0.893	0.8485

Tabla 1: Comparativa de ExPose frente a todos los resultados obtenidos

Para obtener estas métricas, se lleva a cabo el proceso de entrenamiento, lo que conduce a la identificación de los tres mejores modelos entrenados (correspondientes a las tres mejores épocas).

Estos modelos seleccionados se emplean posteriormente en la etapa de pruebas para calcular las métricas más destacadas. Este proceso sigue la metodología que se ha establecido para MHFormer, garantizando así una comparación coherente y fiable en términos de resultados.

5.5.4. Tarea 3- Presentación y análisis de resultados

En esta tarea, se realiza un análisis detallado de los resultados. El objetivo es comprender en profundidad el rendimiento del modelo y las razones detrás de ciertos patrones observados en las métricas de evaluación. Los resultados finales obtenidos se presentan en el *capítulo 7.2 - Análisis de los resultados obtenidos por la integración de ExPose*

5.5.5. Evaluación del sprint

Este último *sprint*, sirve como guinda al trabajo realizado ya que deja relucir el potencial del modelo, así como el impacto del mismo. Se llevan a cabo una serie de tareas clave que no solo consolidan el trabajo previo, sino que también ofrecen una comprensión más profunda y completa de cómo ExPose se integra en el contexto de Bodyflow y cómo enriquece y complementa el panorama general de la librería.

El esfuerzo realizado es de 59 horas sobre las 60 previstas. Tras la revisión de los resultados, el Scrum Master secunda el éxito en la realización de las tareas previstas en el sprint y la finalización del proyecto.

6. Estudio económico

En el ámbito de cualquier proyecto, es esencial llevar a cabo un estudio económico para evaluar su viabilidad, incluyendo aquellos relacionados con la ingeniería del software. Este análisis requiere un nivel de detalle suficiente que permita comprender y estimar los aspectos económicos asociados al proyecto, como los costes materiales, humanos y de desarrollo utilizados. Un enfoque minucioso en la evaluación económica garantiza una visión completa de los aspectos financieros del proyecto.

6.1. Costes económicos

6.1.1. Costes materiales

En lo relativo a los costes materiales, se requiere un ordenador. Aunque las características mínimas con las que debería contar el mismo se establecen en 1GHZ de procesador con dos núcleos, una memoria de 4GB y un almacenamiento de 64GB, se valora el costo de los ordenadores en uso. Primero la utilización de mi ordenador personal durante 6 meses HP ENVY Laptop model13-aq0003ns por valor de 1370,00€, y una vida útil de 5 años valorándose sus costes de amortización asociados en 137,00€. El seleccionado por ITAINNOVA es el HP ProBook 650 G8 Intel Core i5-1135G7/8GB/256GB SSD/15.6". Este ordenador tiene un coste actualmente de 709,00€, y una vida útil estimada de 4 años. Al haberme cedido el ordenador durante 6 meses sus costes se valoran en 88,63€.

En segundo lugar, se utiliza un iPad (8ª gen) para la recopilación de información y realización de la memoria valorado en 379,00€ y 4 años de vida útil. Este es utilizado durante un año así que se valora su costo en 94,75€.

A lo largo del proyecto también se ha hecho uso de otros dispositivos para la realización de pruebas y su monitorización, un iPhone 13 Pro, un Apple Watch Ultra, 2 GoPro HERO11 Black, 3 Xsens DOT entre otros. Estos recursos no se consideran entendiendo que son dispositivos en propiedad de ITAINNOVA, están amortizados y que además dichas pruebas no están exclusivamente dirigidas al desarrollo de este proyecto. Sin embargo, existen unos costes asociados al funcionamiento del material, la electricidad, la proporción de costos de refrigeración, la iluminación de oficinas y la depreciación del equipamiento de las instalaciones, entre otros elementos, que se traducen en un costo fijo calculado por la empresa de 390,00€.

Recurso	Valoración
HP ENVY Laptop model13-aq0003ns	137,00€
HP ProBook	88,63€
iPad (8ª gen)	94,75€
Funcionamiento del material	390,00€
TOTAL	710,38€

Tabla 2: Resumen y total de los costes materiales

6.1.2. Costes humanos

En relación con los costes humanos, la parte más importante es desarrollada por el Equipo de Desarrollo. Si consideramos el puesto de Técnico/a I+D+I en Big Data / Inteligencia Artificial, el cual estaría virtualmente desarrollando en ITAINNOVA, su rango salarial oscila entre 30.285,15 y 43.264,50 euros brutos anuales [30]. Dado que no se cuenta con experiencia previa, utilizaremos el valor más bajo de esta banda salarial.

Por tanto, el coste del empleado de manera mensual es de 3.335,15 €. Considerando 160 horas mensuales de trabajo el coste por hora asciende a los 20,84 €.

Salario bruto	2.523,76 €
Contingencias Comunes	595,61 €
Accidentes de Trabajo y Enfermedades Profesionales (AT y EP)	56,78 €
Desempleo	138,81 €
Formación Profesional	15,14 €
Fondo de garantía salarial (FOGASA)	5,05 €
COSTE TOTAL DEL EMPLEADO	3.335,15 €

Tabla 3: Coste empresa mensual Equipo de Desarrollo

Estimar el tiempo real invertido por otras personas que han podido intervenir en cierto momento del proyecto junto con sus respectivos salarios resulta realmente complicado; sin embargo, incluiremos al Scrum Master en esta estimación para alcanzar una valoración económica lo más precisa posible. La banda salarial en el puesto de Project Manager en Big Data / Inteligencia Artificial que es el que desempeña en este proyecto en ITAINNOVA tiene una banda salarial entre 43.264,50 y 53.758,03 euros. Considerando su experiencia y posición, asumiremos el salario más alto dentro de la banda salarial.

Por tanto, el coste del empleado de manera mensual es de 5.920,10 €. Considerando 160 horas mensuales de trabajo el coste por hora asciende a los 37,00 €.

Salario bruto	4.479,84 €
Contingencias Comunes	1.057,24 €
Accidentes de Trabajo y Enfermedades Profesionales (AT y EP)	100,80 €
Desempleo	246,39 €
Formación Profesional	26,88 €
Fondo de garantía salarial (FOGASA)	8,96 €
COSTE TOTAL DEL EMPLEADO	5.920,10 €

Tabla 4: Coste empresa mensual Scrum Master

En los desgloses de salarios proporcionados, se puede notar la ausencia de los complementos salariales. Esta aparente omisión puede explicarse debido a que las pagas extras son remuneradas de manera prorrateada a lo largo del año. En lugar de recibir una suma considerable en un solo mes, dichos montos se distribuyen de forma equitativa en las nóminas mensuales regulares.

Recurso	Horas	Coste/Hora	Coste
Equipo de desarrollo	338	20,84 €	7.043,92 €
Scrum Master	20	37,00 €	740,00 €
TOTAL			7.783,92 €

Tabla 5: Costes humanos totales en el proyecto

6.1.3. Costes de desarrollo

Finalmente debemos tener en cuenta los costes de desarrollo, que incluyen licencias de software, infraestructuras y formación, dentro del proyecto el peso principal del entorno de desarrollo recae sobre la plataforma Visual Studio Code cuya licencia de uso no genera costes en el proyecto. Sin embargo, el coste computacional en los servidores de cálculo *hypathia* y *hopper* de ITAINNOVA si pueden considerarse importantes ascendiendo a los 112,90 € teniendo en cuenta solo su uso, es decir el costo de energía $(200 W * 430 horas) * €0,15/kWh = €12,90$ y el costo de mantenimiento valorado en 100,00 € de manera prorrateada en horas dirigidas al proyecto. No se consideran los costes de seguridad, licencias o depreciación ya que el uso específico de los servidores para este proyecto representa una carga mínima.

Además, el proyecto requiere de otras herramientas como Microsoft 365 Empresa Estándar con un coste de 11,70 € por usuario al mes (durante 12 meses) [31] también deberíamos considerar el acceso a los *papers*, aunque en este caso se ha realizado de manera gratuita gracias a la licencia de la universidad.

Costes	Valoración
Microsoft 365	140,00€
Computacionales	112,90€
TOTALES	252,90€

Tabla 6: Costes de desarrollo

6.1.4. Costes Totales

Después de considerar todos los costes del proyecto, se elabora una tabla resumen para tener una visión clara y concisa del coste real.

Costes	Valoración
Materiales	710,38 €
Humanos	7.783,92 €
Desarrollo	252,90€
TOTALES	8.747,20€

Tabla 7: Costes totales

6.2. Beneficios económicos

Medir el beneficio económico directo de este proyecto puede ser complicado, ya que se trata de una integración en una librería de propósito investigador y no se enfoca necesariamente en obtener ingresos financieros de manera directa.

Aunque la medición económica directa puede carecer de sentido, es esencial resaltar los beneficios tangibles e intangibles que este proyecto ofrece, tanto para la institución involucrada como para la sociedad en su conjunto. La capacidad de compartir conocimiento a través de la herramienta de código abierto puede tener un impacto significativo en el campo de la investigación y el desarrollo de tecnologías futuras.

7. Resultados

En esta sección, se analiza, en primer lugar, el desarrollo del proyecto y el desempeño obtenido a través de Scrumban, y, en segundo lugar, se presentan y analizan los resultados finales de la aplicación de detección de pose.

7.1. Desarrollo del proyecto con Scrumban

A lo largo del proyecto, ha quedado claro que se han abordado dos aspectos fundamentales: la investigación en el campo de la visión y el desarrollo tecnológico a través de la integración de ExPose, junto con la actualización del módulo de sincronización. Adicionalmente, el proyecto va más allá al culminar con la obtención y comparación de métricas, las cuales no solo validan el enfoque adoptado, sino que también aportan un mayor valor al conjunto del trabajo realizado. En este proyecto, se ha trabajado meticulosamente para fusionar la investigación y el desarrollo tecnológico, generando resultados que destacan tanto la solidez de la propuesta como su impacto en la mejora de la detección de poses y su aplicación en diversos contextos.

En términos de metodología, se ha encontrado que la elección de la metodología Scrumban para el proyecto ha sido muy acertada. Dado el carácter investigador e innovador de ciertos aspectos del proyecto, en ocasiones las estimaciones iniciales de tiempo resultaron ser algo ajustadas. Sin embargo, la flexibilidad inherente a la metodología permitió abordar estos desafíos de manera exitosa. Además, se produjeron incrementos en los objetivos del proyecto en varias instancias, incluyendo una completa integración de un nuevo módulo y corrección de errores, así como la actualización de un módulo de sincronización obsoleto. La gestión de estos ajustes fue posible de manera efectiva gracias a los fundamentos ágiles de Scrumban.

Se ha trabajado en función de los objetivos establecidos, los cuales se han desglosado en tareas asignadas a cada sprint. Al utilizar la metodología Scrumban, no se ha elaborado un gráfico inicial de "cómo vamos a trabajar a lo largo del proyecto", ya que se ha priorizado la visualización de los esfuerzos restantes, lo cual proporciona mayor claridad y valor durante la ejecución. Sin embargo, en la actualidad se reconoce la importancia de los datos recopilados para mejorar la planificación de proyectos en retrospectiva. Por lo tanto, se presenta un gráfico que ilustra la evolución del trabajo en función de los objetivos a lo largo del tiempo en la *Figura 37*.

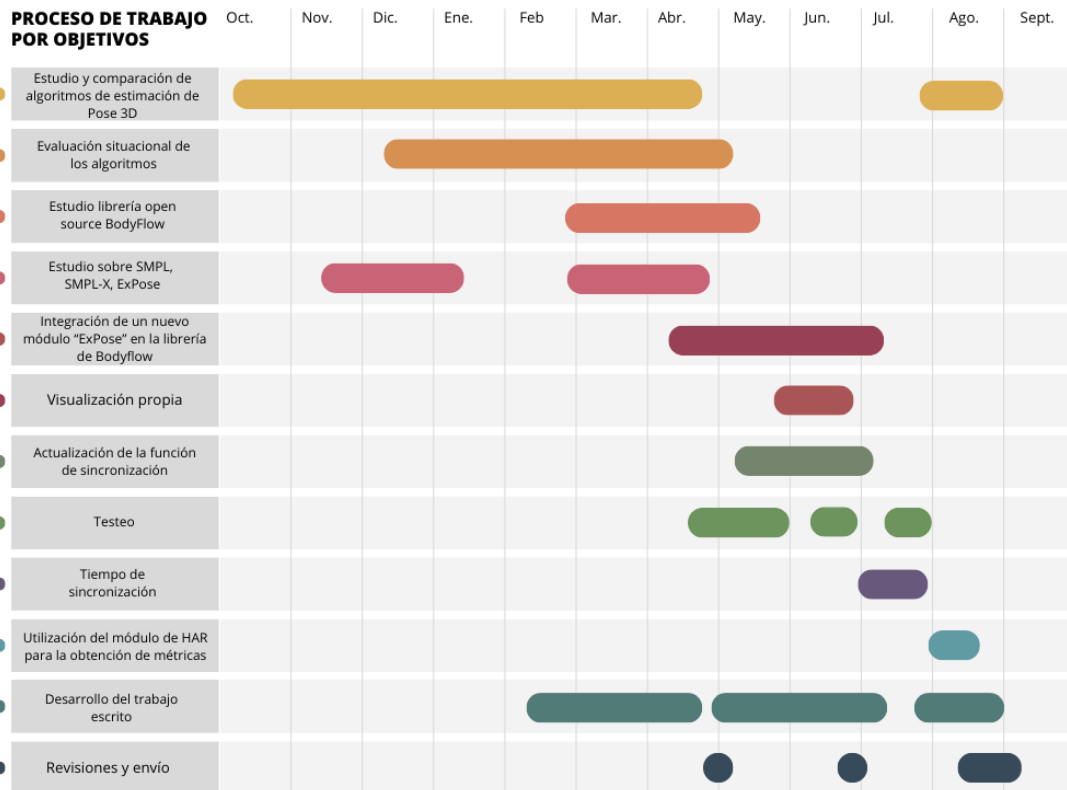


Figura 37: Gráfica retrospectiva del progreso del proyecto, que ilustra cómo se ha desarrollado a lo largo del año natural. (Elaboración propia)

Aunque el proyecto engloba prácticamente un año natural es importante destacar que los esfuerzos invertidos no han sido constantes en todos los meses, sino que han variado en función de las fases del proyecto y las tareas específicas que se abordaban en cada periodo.

La estimación y el esfuerzo real que se ha mencionado en cada sprint, ahora deben ser representados de manera conjunta para ver como finalmente se ha conseguido completar el proyecto bajo estimación. Con el fin de alcanzar este objetivo, se emplea un diagrama de Gantt. (ver Figura 38)

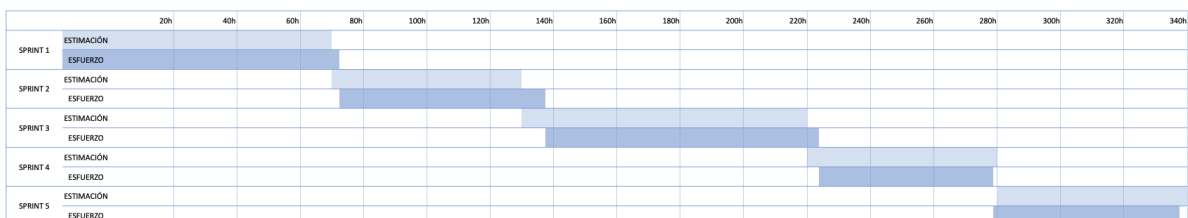


Figura 38: Diagrama de Gantt: En el gráfico, se pueden apreciar en tono azul claro las estimaciones que se han efectuado sprint tras sprint, mientras que en azul oscuro se representa el esfuerzo efectivo empleado. (Elaboración propia)

Por último, se presenta la *Tabla 8*, que muestra el desarrollo de los *sprints* en términos de fechas y objetivos alcanzados. Es importante señalar que en el *sprint* 3 se tuvo que retomar una tarea del *sprint* anterior, y se destaca que el esfuerzo total realizado es ligeramente inferior a las estimaciones realizadas *sprint* tras *sprint*. Las estimaciones consideran el tiempo asignado para la redacción de la memoria, que totaliza 80 horas.







	TITULO	ESTIMACIÓN	ESFUERZO	FECHA INICIO	FECHA FIN	OBJETIVOS ALCANZADOS	REVISIÓN SCRUM MASTER
1	Estimación de Pose, ExPose y Bodyflow	70 horas	72 horas	14/10/22	18/01/23	100%	
2	Preparación del entorno y análisis del estado actual	60 horas	65 horas	19/01/23	16/04/23	92%	
3	Integración de ExPose y visualización de resultados	90 horas	86 horas	17/04/23	25/05/23	100%	
4	Función de sincronización	60 horas	56 horas	26/05/23	03/07/23	100%	
5	Evaluación del modelo y resultados	60 horas	59 horas	04/07/23	30/08/23	100%	
		340 horas	338 horas	14/10/22	30/08/23	100%	

Tabla 8: Resumen de los sprints

Indudablemente, desde una perspectiva profesional, podemos considerar este proyecto como un rotundo éxito. Se han cumplido y superado todos los objetivos establecidos en un plazo menor al estimado. Además, el resultado final es un producto innovador y mejorado que se integra de manera fluida con las soluciones previas, aportando un nuevo nivel de funcionalidad y rendimiento. Esta culminación exitosa valida no solo la visión y el enfoque del proyecto, sino también el compromiso y el arduo trabajo del proyectando involucrado en su desarrollo. Sin duda, este logro demuestra la capacidad de enfrentar desafíos técnicos y cumplir con las expectativas, estableciendo un nuevo estándar en la detección y seguimiento de movimientos humanos en entornos tridimensionales.

7.2. Análisis de los resultados obtenidos por la integración de ExPose

En este apartado se presenta el análisis llevado a cabo en la tarea 3 del quinto *sprint*, en la cual se emplean los resultados obtenidos para evaluar el rendimiento y funcionamiento del módulo que ha sido integrado.

Es evidente observando a simple vista la *Tabla 1* (ver página 57) que, los resultados logrados por ExPose son notoriamente valiosos. Sus métricas superan a las obtenidas por los sensores y se asemejan en gran medida, a las de MHFormer. De hecho, si nos centramos solo en el valor de la exactitud (*accuracy*) ExPose se elige como superior como podemos observar en la *Tabla 1.1*

HAR Model	Features	F1	Accuracy	Precision
Transformer	all (ExPose)	0.803	0.944	0.836
LSTM	all (ExPose)	0.781	0.937	0.797
CNN	all (ExPose)	0.761	0.937	0.763
Transformer	3d (ExPose)	0.836	0.931	0.827
LSTM	imus	0.702	0.916	0.718
CNN	3d (ExPose)	0.817	0.913	0.833
Transformer	imus	0.793	0.896	0.852
LSTM	3d (ExPose)	0.746	0.893	0.767
CNN	all (MHFormer)	0.781	0.892	0.816
Transformer	all (MHFormer)	0.865	0.891	0.882
LSTM	3d (MHFormer)	0.787	0.888	0.795
Transformer	3d (MHFormer)	0.825	0.881	0.818
Transformer	2d	0.825	0.864	0.819
LSTM	2d	0.692	0.864	0.702
LSTM	all (MHFormer)	0.718	0.848	0.738
CNN	3d (MHFormer)	0.760	0.836	0.797
CNN	2d	0.616	0.835	0.676
CNN	imus	0.655	0.833	0.685
CNN	ankle	0.664	0.795	0.699
Transformer	ankle	0.651	0.785	0.679
LSTM	ankle	0.515	0.768	0.550

Tabla 1.1: Comparativa de resultados ordenados por Accuracy

Por otro lado, al enfocarnos exclusivamente en los datos generados por ExPose (ver *Tabla 1.2*), se hace evidente que siempre se obtienen mejores resultados de exactitud con la característica "all". Esto resalta claramente la ventaja de contar con sensores además de los datos capturados únicamente por cámaras. Las demás métricas no presentan un patrón discernible debido a su similitud. Sin embargo, es posible notar que la utilización del modelo Transformer logra mejores resultados.

HAR Model	Features	F1	Accuracy	Precision
Transformer	all (ExPose)	0.803	0.944	0.836
LSTM	all (ExPose)	0.781	0.937	0.797
CNN	all (ExPose)	0.761	0.937	0.763
Transformer	3d (ExPose)	0.836	0.931	0.827
CNN	3d (ExPose)	0.817	0.913	0.833
LSTM	3d (ExPose)	0.746	0.893	0.767

Tabla 1.2: Comparativa de los resultados de ExPose ordenados por Accuracy

Si bien las conclusiones obtenidas hasta este punto son notablemente positivas, es importante resaltar que ExPose no se basa en una predicción 2D, es decir, no utiliza dicha predicción en su enfoque. Por lo tanto, la comparación más relevante se establece entre los resultados de ExPose y los resultados de MHFormer + CPN, ya que este último utiliza los keypoints 2D en su estimación de pose. Estos resultados son los marcados como 3d (ExPose) y 3d (MHFormer) respectivamente.

HAR Model	Features	F1	Accuracy	Precision
Transformer	3d (ExPose)	0.836	0.931	0.827
Transformer	3d (MHFormer)	0.825	0.881	0.818
CNN	3d (ExPose)	0.817	0.913	0.833
LSTM	3d (MHFormer)	0.787	0.888	0.795
CNN	3d (MHFormer)	0.760	0.836	0.797
LSTM	3d (ExPose)	0.746	0.893	0.767

Tabla 1.3: Comparativa resultados 3d ordenados por F1

HAR Model	Features	F1	Accuracy	Precision
Transformer	3d (ExPose)	0.836	0.931	0.827
CNN	3d (ExPose)	0.817	0.913	0.833
LSTM	3d (ExPose)	0.746	0.893	0.767
LSTM	3d (MHFormer)	0.787	0.888	0.795
Transformer	3d (MHFormer)	0.825	0.881	0.818
CNN	3d (MHFormer)	0.760	0.836	0.797

Tabla 1.4: Comparativa resultados 3d ordenados por Accuracy (Exactitud)

HAR Model	Features	F1	Accuracy	Precision
CNN	3d (ExPose)	0.817	0.913	0.833
Transformer	3d (ExPose)	0.836	0.931	0.827
Transformer	3d (MHFormer)	0.825	0.881	0.818
CNN	3d (MHFormer)	0.760	0.836	0.797
LSTM	3d (MHFormer)	0.787	0.888	0.795
LSTM	3d (ExPose)	0.746	0.893	0.767

Tabla 1.5: Comparativa resultados 3d ordenados por Precisión

Tal como se puede observar en las *Tablas 1.3, 1.4 y 1.5*, independientemente de la métrica utilizada, el modelo ExPose en una de sus combinaciones con un modelo HAR siempre presenta los resultados más destacados. Estos resultados adquieren una gran relevancia, ya que ExPose no solo ofrece una mejor detección de la pose sin depender de un modelo 2D para su obtención, sino que también agrega un valor añadido al proporcionar la malla 3D.

De cualquier forma, podemos tener una visión más clara a través de las matrices de confusión, estas tienen una estructura de filas y columnas, donde cada fila representa la clase real y cada columna representa la clase predicha por el modelo. Los elementos en la diagonal principal representan las predicciones correctas, mientras que los elementos fuera de la diagonal principal representan las predicciones incorrectas.

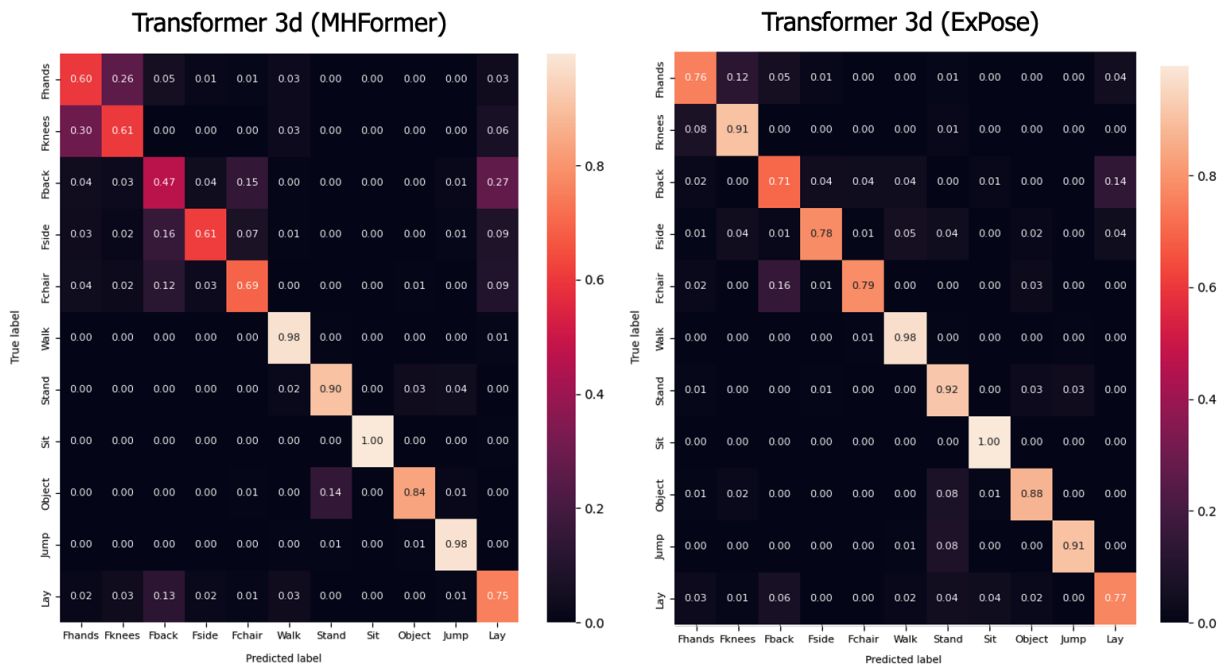


Figura 39: Matrices de confusión obtenidas por MHFormer y ExPose junto con el modelo Transformer de HAR

Resulta evidente la eficacia de ExPose en los resultados obtenidos por el modelo. Los errores principales de ExPose centran en situaciones concretas en las que, por ejemplo, cuando una persona cae de una silla, el modelo predice en un 16% de los casos que se trata de una caída hacia atrás, y cuando efectivamente ocurre una caída hacia atrás, el modelo predice en un 14% que la persona está tumbada.

Si profundizamos más, en situaciones donde una persona está agarrando un objeto o realizando un salto, en un 8% de las veces se predice que la persona está de pie. En resumen, los errores

más recurrentes parecen tener una lógica subyacente, lo que indica que el modelo está cometiendo fallos tan solo en contextos particulares.

En resumen, los resultados obtenidos enfatizan la robustez de ExPose en cuanto a la detección y representación de poses humanas en 3D. Su capacidad para funcionar sin depender de predicciones en 2D es un distintivo que le otorga una ventaja significativa. Estas métricas no solo enriquecen el trabajo desarrollado mediante la integración, sino que también resaltan el potencial que puede tener en aplicaciones de detección y seguimiento de movimientos humanos.

8. Conclusiones

Los objetivos delineados en la propuesta se han completado de manera exitosa. Se ha llevado a cabo un análisis profundo sobre algoritmos de estimación de pose, así como un minucioso escrutinio de la librería BodyFlow, destacando la definición tanto de sus características actuales como de las venideras. Adicionalmente, se ha realizado una investigación pormenorizada de los modelos SMPL, SMPL-X y ExPose, centrandó la atención en su capacidad para generar resultados en la detección de pose. Para redondear el proyecto se ha aplicado el conocimiento de manera práctica.

En resumen, este proyecto ha abordado de manera integral la integración y análisis de ExPose en el entorno de BodyFlow, demostrando su eficacia en la detección y representación de poses humanas en 3D a partir de imágenes RGB. A través de un enfoque metódico y detallado, se ha logrado adaptar ExPose para su utilización, optimizando su rendimiento y asegurando su funcionalidad dentro de la librería.

La integración implicó una cuidadosa planificación y ejecución de distintas fases, desde la adaptación de licencias y códigos hasta la creación de un *wrapper* y la generación de resultados visuales y funcionales. Los resultados obtenidos de las evaluaciones y pruebas demuestran que ExPose es capaz de generar representaciones precisas y detalladas de poses humanas en 3D, superando incluso a los sensores y presentando métricas comparables a las de MHFormer. Esto refuerza la efectividad y la versatilidad de ExPose en diversas aplicaciones que involucran la detección y representación de movimientos humanos.

Los desafíos principales que se han enfrentado son la coordinación con las actualizaciones en la librería, ya que se estaban introduciendo nuevas funcionalidades y se estaba llevando a cabo la puesta en producción de la misma en paralelo con el desarrollo del PFG. Además, se ha tenido que abordar el problema del costo computacional asociado a la ejecución de las pruebas y al procesamiento del conjunto de datos utilizado para obtener los resultados finales. Por otro lado, las limitaciones se han reducido a dos factores clave, el tiempo efectivo disponible para la realización del PFG siendo crucial que el alcance del proyecto fuera realista y acorde con los recursos disponibles y la capacidad de implementación, dado que la solución óptima no siempre es factible de implementar debido a restricciones legales y técnicas.

En última instancia, es esencial destacar que, si bien los logros obtenidos con ExPose son notables, no podemos obviar el costo de desarrollo asociado. A lo largo del proyecto, quedó claro que la integración y optimización de ExPose así como la refactorización del módulo de sincronización demandaron una inversión considerable en términos de tiempo y recursos. A pesar de este desafío, los resultados obtenidos y la mejora en las capacidades de detección y representación de poses en 3D refuerzan la importancia de estos esfuerzos. Esta inversión no

solo ha impulsado el rendimiento del modelo, sino que también ha establecido las bases para un futuro desarrollo en la detección y seguimiento de movimientos humanos en una variedad de aplicaciones, que abarcan desde la supervisión del rendimiento físico y el bienestar hasta la seguridad y vigilancia, la automatización industrial, e incluso el entretenimiento en videojuegos, demostrando su impacto en diversos campos.

8.1. Extensiones y mejoras futuras

A medida que este proyecto llega a su conclusión, es importante destacar que el potencial de desarrollo y aplicación de Bodyflow es vasto y lleno de oportunidades. Si bien las posibilidades son prácticamente infinitas, en esta sección nos centraremos en las direcciones que se perfilan como inminentes en el ámbito de ITAINNOVA. Estas direcciones representan áreas concretas en las que Bodyflow está listo para marcar un impacto significativo y ofrecer soluciones innovadoras. A continuación, exploraremos algunas de las direcciones más prometedoras que están en el horizonte de ITAINNOVA y que podrían aprovechar al máximo la versatilidad y adaptabilidad de Bodyflow.

- a. Exploración del Tercer Módulo- Deterioro cognitivo. Además de los dos módulos ya implementados, existe el potencial para desarrollar un tercer módulo que se focalice en la detección de deterioros cognitivos. Este nuevo enfoque podría aprovechar las capacidades y flexibilidad de Bodyflow para analizar imágenes médicas y detectar posibles anomalías asociadas. Este sería un paso significativo hacia la aplicación de la tecnología en campos médicos de alta relevancia.
- b. Colaboración con Hospitales y Empresas Sanitarias. Una dirección prometedora es la colaboración con hospitales y organizaciones para obtener resultados concretos en entornos clínicos. La aplicación de Bodyflow en colaboración con profesionales de la salud podría aportar información valiosa para la detección y el análisis de patologías, mejorando así la toma de decisiones médicas.
- c. Resultados y Aplicaciones en el Ámbito Deportivo. Otro camino es la obtención de resultados y aplicaciones en el deporte en colaboración con algún equipo. Utilizar Bodyflow para evaluar la biomecánica y el rendimiento de los jugadores podría proporcionar *insights* únicos para el entrenamiento y la prevención de lesiones, beneficiando tanto a los deportistas como a los equipos.
- d. Continua Investigación y Comparación de Métodos. Con la implementación exitosa de ExPose, se abre la puerta para seguir investigando y comparando diferentes métodos de detección de pose. Dado que Bodyflow es una librería abierta y modular, se podrían incorporar y evaluar otros enfoques para determinar cuál ofrece el mejor rendimiento en diferentes contextos.

9. Bibliografía

- [1] Marina Azpiroz, Rafael del Hoyo Alonso, Irene López Bosque, Carlos Marañes, Ángel Gimeno Valero, Pilar Salvo Ibáñez, José Luis Pérez. PT3. SARCOPENIA. ANÁLISIS Y DISEÑO. Diciembre 2022
- [2] Joan Reig Doménech. Estudio del estado del arte de los métodos de estimación de la pose humana en 3D. (Junio 2018). Disponible en: <https://core.ac.uk/download/pdf/159143016.pdf>
- [3] Wu Liu, Member, IEEE, Qian Bao, Yu Sun, and Tao Mei, Fellow. Recent Advances in Monocular 2D and 3D Human Pose Estimation: A Deep Learning Perspective., IEEE. JOURNAL OF LATEX CLASS FILES VOL. X, NO. X, (2021). Disponible en: <https://arxiv.org/pdf/2104.11536.pdf>
- [4] Yating Tian, Hongwen Zhang, Yebin Liu, Member, IEEE, and Limin Wang Recovering 3D Human Mesh from Monocular Images: A Survey. IEEE. arXiv:2203.01923v2 [cs.CV] (8 marzo 2022). Disponible en: <https://arxiv.org/pdf/2203.01923.pdf>
- [5] Nikolaos Sarafianos, Bogdan Boteanub, Bogdan Ionescub, Ioannis A. Kakadiaris. 3D Human pose estimation: A review of the literature and analysis of covariates. (8 septiembre 2016)
- [6] Gavrilu, D. M. 'The Visual Analysis of Human Movement: A Survey', Computer Vision and Image Understanding (1999)
- [7] Moeslund, T. B., Hilton, A. and Krüger, V. A survey of advances in vision-based human motion capture and analysis, Computer Vision and Image Understanding (2006)
- [8] Poppe, R. Vision-based human motion analysis: An overview, Computer Vision and Image Understanding (2007)
- [9] Kondoři, F. A. Bring Your Body into Action (2014). Disponible en: <http://umu.divaportal.org/smash/get/diva2:716122/FULLTEXT01.pdf>.
- [10] Suma, E. A. et al. FAAST: The flexible action and articulated skeleton toolkit, Proceedings - IEEE Virtual Reality (2011)
- [11] Condell, J., Moore, G. and Moore, J. Software and methods for motion capture and tracking in animation. The 2006 International Conference on Computer Graphics and Virtual Reality.
- [12] Ortiz-catalan, M. et al. Emerging therapies in Neurorehabilitation, Emerging Therapies in Neurorehabilitation Biosystems & Biorobotics Volume 4. (2014)
- [13] Banerjee, Prithviraj & Sengupta, Somnath. Human motion detection and tracking for video surveillance. (Enero 2008). Disponible en: https://www.researchgate.net/publication/228861429_Human_motion_detection_and_tracking_for_video_surveillance

-
- [14] A. O. Balan, L. Sigal, M. J. Black, J. E. Davis, and H. W. Haussecker Detailed human shape and pose from images, in CVPR, (2007).
- [15] D. Anguelov, P. Srinivasan, D. Koller, S. Thrun, J. Rodgers, and J. Davis. SCAPE: shape completion and animation of people. TOG, (2005).
- [16] M. Loper, N. Mahmood, J. Romero, G. Pons-Moll, and M. J. Black. Smpl: A skinned multi-person linear model. TOG, (2015). Disponible en: <https://files.is.tue.mpg.de/black/papers/SMPL2015.pdf>
- [17] F. Bogo, A. Kanazawa, C. Lassner, P. Gehler, J. Romero, and M. J. Black. Keep it smpl: Automatic estimation of 3d human pose and shape from a single image , in ECCV, (2016). Disponible en: <https://files.is.tue.mpg.de/black/papers/BogoECCV2016.pdf>
- [18] M. Loper, N. Mahmood, and M. J. Black. TOG, (2014). Mosh: Motion and shape capture from sparse markers. Disponible en: <https://files.is.tue.mpg.de/black/papers/MoSh.pdf>
- [19] C. Lassner, J. Romero, M. Kiefel, F. Bogo, M. J. Black, and P. V. Gehler. Unite the people: Closing the loop between 3d and 2d human representations. CVPR, (2017).
- [20] C. Ionescu, D. Papava, V. Olaru, and C. Sminchisescu. Human3. 6m: Large scale datasets and predictive methods for 3d human sensing in natural environments. TPAMI, (2014)
- [21] N. Mahmood, N. Ghorbani, N. F. Troje, G. Pons-Moll, and M. J. Black. Amass: Archive of motion capture as surface shapes. ICCV (2019).
- [22] G. Pavlakos, V. Choutas, N. Ghorbani, T. Bolkart, A. A. Osman, D. Tzionas, and M. J. Black. Expressive body capture: 3d hands, face, and body from a single image . CVPR, (2019). Disponible en: https://ps.is.mpg.de/uploads_file/attachment/attachment/497/SMPL-X.pdf
- [23] M. Kocabas, N. Athanasiou, and M. J. Black. Vibe: Video inference for human body pose and shape estimation, CVPR, (2020).
- [24] Jindong Wang, Yiqiang Chen, Shuji Hao, Xiaohui Peng, Lisha Hu. Deep learning for sensor-based activity recognition: A Survey. Pattern Recognition Letters, (2019). Disponible en: https://jd92.wang/assets/files/a10_prl18.pdf
- [25] Arshad MH, Bilal M, Gani A. Human Activity Recognition: Review, Taxonomy and Open Challenges, Sensors, 2022. Disponible en: <https://www.mdpi.com/1424-8220/22/17/6463>
- [26] Julio Roche. Scrumban, la metodología del cambio. Disponible en: <https://www2.deloitte.com/es/es/pages/technology/articles/scrumban.html>
- [27] Julia Martins. Scrum: conceptos clave y cómo se aplica en la gestión de proyectos (9 de junio de 2023) Disponible en: <https://asana.com/es/resources/what-is-scrum>
- [28] Max Rehkopf. Sprints de scrum. Disponible en: <https://www.atlassian.com/es/agile/scrum/sprints>
- [29] L. Martínez-Villaseñor, H. Ponce, J. Brieva, E. Moya-Albor, J. Núñez-Martínez, and C. Peñafort-Asturiano, "Up-fall detection dataset: A multimodal approach," Sensors,

-
- vol. 19, p. 1988, Apr 2019.
- [30] CATALOGO PERSONAL LP ITA_TRANSPARENCIA. Disponible en:
https://transparencia.aragon.es/sites/default/files/documents/rpt_ita.pdf
- [31] Microsoft Planes y Precios. Disponible en: <https://www.microsoft.com/es-es/microsoft-365/business/compare-all-microsoft-365-business-products#>
- [32] Redes Neuronales: Conceptos Básicos y Aplicaciones. Damián Jorge Matich. Informática Aplicada a la Ingeniería de Procesos – Orientación I
- [33] CONCEPTOS BÁSICOS SOBRE REDES NEURONALES. Disponible en:
<http://grupo.us.es/gtocom/pid/pid10/RedesNeuronales.htm>
- [34] Learning and Neural Network classifiers. Rafael del Hoyo Alonso. Intelligent Systems, ITAINNOVA
- [35] SMPL: A Skinned Multi-Person Linear Model. Michael Black. Video presented in (SIGGRAPH Asia 2015) <https://www.youtube.com/watch?v=kuBIUyHeV5U&t=11s>
- [36] SMPLify. Michael Black (ECCV 2016, Supplementary Material)
<https://www.youtube.com/watch?v=OgX49T2Cqdo>
- [37] Código abierto simplify. Disponible en: <https://github.com/vchoutas/smplify-x>
- [38] CONDA CHEAT SHEET Disponible en:
<https://docs.conda.io/projects/conda/en/4.6.0/downloads/52a95608c49671267e40c689e0bc00ca/conda-cheatsheet.pdf>
- [39] User Guide Conda Disponible en: <https://docs.conda.io/projects/conda/en/latest/user-guide/getting-started.html>
- [40] Math Works. What Is Object Detection? Disponible en:
<https://es.mathworks.com/discovery/object-detection.html>
- [41] Vasileios Choutas, Georgios Pavlakos, Timo Bolkart, Dimitrios Tzionas and Michael J. Black. Monocular Expressive Body Regression through Body-Driven Attention. Disponible en: https://ps.is.mpg.de/uploads_file/attachment/attachment/620/0983.pdf
- [42] Vasileios Choutas, Georgios Pavlakos, Timo Bolkart, Dimitrios Tzionas and Michael J. Black. Monocular Expressive Body Regression through Body-Driven Attention
Supplemental Material Disponible en:
https://ps.is.mpg.de/uploads_file/attachment/attachment/621/0983-supp_no_header_compressed.pdf
- [43] Pixel Fondue. How It Works | Linear Blend Skinning . Disponible en:
https://www.youtube.com/watch?v=QDXG4wNzkOE&ab_channel=PixelFondue
- [44] Irene López-Bosque, Carlos Marañes-Nueno, Ana Caren Hernández-Ruiz, Rocío Aznar Gimeno, Pilar Salvo-Ibañez, M^a de la Vega Rodrigálvarez-Chamarro, David Abadía Gallego, Rafael del-Hoyo Alonso. Bodyflow: A library for human pose estimation and activity recognition
-

-
- [45] R. Espinosa, H. Ponce, S. Gutiérrez, L. Martínez-Villaseñor, J. Brieva, and E. Moya-Albor, "A vision-based approach for fall detection using multiple cameras and convolutional neural networks: A case study using the up-fall detection dataset," *Computers in Biology and Medicine*, vol. 115, p. 103520, 2019.
- [46] J. J. P. Suarez, N. Orillaza, and P. Naval, "Afar: A real-time vision-based activity monitoring and fall detection framework using 1d convolutional neural networks," in *2022 14th International Conference on Machine Learning and Computing (ICMLC), ICMLC 2022, (New York, NY, USA)*, p. 555–559, Association for Computing Machinery, 2022.
- [47] Manon Kok, Jeroen D. Hol and Thomas B. Schön (2017), "Using Inertial Sensors for Position and Orientation Estimation", *Foundations and Trends® in Signal Processing: Vol. 11: No. 1-2*, pp 1-153. Disponible en: <https://www.nowpublishers.com/article/Details/SIG-094>
- [48] Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation". (2014) Disponible en: <https://arxiv.org/abs/1311.2524>

ANEXO I: Propuesta de Proyecto

Nombre alumno: Marcos Marina Castelló

Titulación: Ingeniería Informática

Curso académico: 2022/23

1. TÍTULO DEL PROYECTO

Desarrollo de un sistema de reconocimiento de poses 3D en personas utilizando ML

2. DESCRIPCIÓN Y JUSTIFICACIÓN DEL TEMA A TRATAR

La detección de pose humana es una tarea bastante relevante en el campo de la visión por computador, que consiste en identificar la postura de una figura humana a partir de una imagen. Dicha pose se define a partir de una serie de puntos clave, que habitualmente serán articulaciones, de forma que el objetivo será encontrar la posición (x, y) de cada uno de esos puntos, o (x, y, z) en el caso de que queramos identificar también la profundidad.

La identificación de pose permite que los usuarios de sistemas de realidad virtual y aumentada puedan interactuar con ellos de forma más natural, así como también se pueden utilizar para identificar eventos o determinadas actividades en vídeos, para temas clínicos o también deportivos.

3. OBJETIVOS DEL PROYECTO

En este trabajo, se desarrollará una metodología para la estimación de pose 3D a partir de una representación reducida de manera robusta y rápida mediante algoritmos de ML. Se estudiarán los algoritmos de estimación de pose 3D más utilizados, determinando sus ventajas y desventajas, así como su representación en un base de datos. Se probarán estos algoritmos de Inteligencia Artificial en diferentes situaciones. Las poses que reconocerá la aplicación se definirán en función del contexto final de uso (como detección de caídas, actividades o eventos que le puedan suceder a una persona).

4. METODOLOGÍA

Se plantean el uso de metodologías ágiles para el desarrollo del proyecto.

5. PLANIFICACIÓN DE TAREAS

Se planificarán reuniones durante el proyecto.

6. OBSERVACIONES ADICIONALES

Tutores académicos: Violeta Monasterio Bazán y Alfonso Santolaria Lorenzo

Tutores profesionales: Vega Rodrigálvarez Chamarro y Rafael del Hoyo Alonso

ANEXO II: Actas de reunión

Reunión 1

REUNIÓN: Primera Reunión

Fecha: 07/09/2022	
Hora comienzo: 11:00	Hora finalización: 12:30
Lugar: ITAINNOVA	
Elabora acta: Marcos Marina	
Convocados: Marcos Marina, Rafael del Hoyo	

Orden del día / Acta

No.	Asunto	Acuerdo
1	Disculpas por ausencia No procede	
2	Aprobar última acta No procede	
3	Asuntos pendientes última acta No procede	
4	Presentación de proyectos por parte del tutor Rafael del Hoyo expone una variedad de proyectos en curso en ITAINNOVA que podrían servir como base para el Proyecto Fin de Grado	001
5	Otros asuntos Se trata la posibilidad de iniciar una colaboración con ITAINNOVA	
6	Próxima reunión Tras la elección del proyecto	

Resumen de acuerdos

Número	Acuerdo	Plazo	Responsable
001	Seleccionar uno de los proyectos propuestos	10/10/2022	Marcos Marina

Reunión 2

REUNIÓN: Entrega solicitud PFG, inicio del proyecto

Fecha: 14/10/2022	
Hora comienzo: 11:15	Hora finalización: 12:00
Lugar: Teams	
Elabora acta: Marcos Marina	
Convocados: Marcos Marina, Rafael del Hoyo	

Orden del día / Acta

No.	Asunto	Acuerdo
1	Disculpas por ausencia No procede	
2	Aprobar última acta No procede	
3	Asuntos pendientes última acta El acuerdo 001 de la anterior acta se da por concluido, tras seleccionar uno de los proyectos propuestos	
4	Revisión de la propuesta de PFG Rafael del Hoyo expone la propuesta tras la selección realizada, se consensúan los objetivos	
5	Bases de conocimiento Rafael del Hoyo le facilita recursos al proyectando para empezar la investigación	001
6	Otros asuntos No se tratan otros asuntos	
7	Próxima reunión 25 de noviembre, para actualizar la situación	

Resumen de acuerdos

Número	Acuerdo	Plazo	Responsable
001	Lectura y comprensión de los <i>papers</i>	25/11/2022	Marcos Marina

Reunión 3

REUNIÓN: Actualización y formación en el proceso investigador

Fecha: 25/11/2022	
Hora comienzo: 20:00	Hora finalización: 21:00
Lugar: Teams	
Elabora acta: Marcos Marina	
Convocados: Marcos Marina, Rafael del Hoyo	

Orden del día / Acta

No.	Asunto	Acuerdo
1	Disculpas por ausencia No procede	
2	Aprobar última acta No procede	
3	Asuntos pendientes última acta El acuerdo 001 de la anterior acta se da por concluido, tras la lectura y comprensión de los <i>papers</i>	
4	Actualización por parte del proyectando Marcos Marina detalla los avances y las impresiones obtenidas en el ámbito de investigación	
5	Formación Rafael del Hoyo realiza una pequeña formación en el ámbito	
6	Otros asuntos No se tratan otros asuntos	
7	Próxima reunión Al completar todas las tareas del <i>sprint</i>	001

Resumen de acuerdos

Número	Acuerdo	Plazo	Responsable
001	Realización de las tareas	18/01/2023	Marcos Marina

Reunión 4

REUNIÓN: Desarrollo de la memoria

Fecha: 08/12/2022	
Hora comienzo: 17:00	Hora finalización: 18:00
Lugar: Teams	
Elabora acta: Marcos Marina	
Convocados: Marcos Marina, Rafael del Hoyo, Violeta Monasterio	

Orden del día / Acta

No.	Asunto	Acuerdo
1	Disculpas por ausencia No procede	
2	Aprobar última acta No procede	
3	Asuntos pendientes última acta No procede	
4	Actualización por parte del proyectando Marcos Marina detalla los avances en el desarrollo de la memoria	
5	Acuerdos verbales Se establecen acuerdos verbales para llevar a cabo la redacción de la memoria en paralelo con la ejecución del proyecto. Además, Marcos Marina presentará los avances en las distintas secciones a Violeta Monasterio.	
6	Otros asuntos No se tratan otros asuntos	
7	Próxima reunión No se especifica	

Resumen de acuerdos

Número	Acuerdo	Plazo	Responsable
-	-	-	-

Reunión 5

REUNIÓN: Fin *sprint* 1. Inicio *sprint* 2

Fecha: 18/01/2023	
Hora comienzo: 15:00	Hora finalización: 16:00
Lugar: Teams	
Elabora acta: Marcos Marina	
Convocados: Marcos Marina, Rafael del Hoyo	

Orden del día / Acta

No.	Asunto	Acuerdo
1	Disculpas por ausencia No procede	
2	Aprobar última acta No procede	
3	Asuntos pendientes última acta El acuerdo 001 del acta de la reunión 3 se da por concluido, tras la realización de las tareas asignadas	
4	Tareas <i>Sprint</i> 2 Se presentan las tareas del segundo <i>sprint</i>	001
5	Otros asuntos Contrato con ITAINNOVA	002
6	Próxima reunión Al completar todas las tareas del <i>sprint</i>	

Resumen de acuerdos

Número	Acuerdo	Plazo	Responsable
001	Realización de las tareas del <i>sprint</i> 2	17/04/2023	Marcos Marina
002	Avance del acuerdo con ITAINNOVA	17/04/2023	Rafael del Hoyo

Reunión 6

REUNIÓN: Fin *sprint* 2. Inicio *sprint* 3

Fecha: 17/04/2023	
Hora comienzo: 18:15	Hora finalización: 19:30
Lugar: Teams	
Elabora acta: Marcos Marina	
Convocados: Marcos Marina, Rafael del Hoyo	

Orden del día / Acta

No.	Asunto	Acuerdo
1	Disculpas por ausencia No procede	
2	Aprobar última acta No procede	
3	Asuntos pendientes última acta El acuerdo 001 de la anterior acta se da por concluido, tras la realización de las tareas asignadas. El acuerdo 002 del acta anterior también se da por concluido al entrar el proyectando en el proceso de selección de ITAINNOVA	
4	Tareas Sprint 3 Se presentan las tareas del tercer <i>sprint</i>	001
5	Otros asuntos No se tratan otros asuntos	
6	Próxima reunión Al completar todas las tareas del <i>sprint</i>	

Resumen de acuerdos

Número	Acuerdo	Plazo	Responsable
001	Realización de las tareas del <i>sprint</i> 3	24/05/2023	Marcos Marina

Reunión 7

REUNIÓN: Actualización del proyecto tras el contrato con ITAINNOVA

Fecha: 28/04/2023	
Hora comienzo: 18:00	Hora finalización: 19:30
Lugar: Teams	
Elabora acta: Marcos Marina	
Convocados: Marcos Marina, Rafael del Hoyo, Violeta Monasterio, Alfonso Santolaria	

Orden del día / Acta

No.	Asunto	Acuerdo
1	Disculpas por ausencia No procede	
2	Aprobar última acta No procede	
3	Asuntos pendientes última acta No procede	
4	Actualización del proyecto Marcos Marina presenta el proyecto a Alfonso Santolaria, Rafael del Hoyo completa la explicación. Violeta ofrece los detalles del avance de la memoria.	
5	Rol de Alfonso Santolaria Se ha acordado que Alfonso desempeñe un papel central al finalizar el proyecto. Una vez que la memoria esté completamente elaborada, se aprovechará la experiencia empresarial de Alfonso para obtener sus opiniones y sugerencias. Esto permitirá a Marcos llevar a cabo las últimas modificaciones necesarias.	001
6	Otros asuntos No se tratan otros asuntos	
7	Próxima reunión No se especifica	

Resumen de acuerdos

Número	Acuerdo	Plazo	Responsables
001	Realización de la versión preliminar de la memoria	20/08/2023	Marcos Marina, Violeta Monasterio

Reunión 8

REUNIÓN: Fin *sprint* 3. Inicio *sprint* 4

Fecha: 25/05/2023	
Hora comienzo: 8:30	Hora finalización: 9:30
Lugar: ITAINNOVA	
Elabora acta: Marcos Marina	
Convocados: Marcos Marina, Rafael del Hoyo	

Orden del día / Acta

No.	Asunto	Acuerdo
1	Disculpas por ausencia No procede	
2	Aprobar última acta No procede	
3	Asuntos pendientes última acta El acuerdo 001 del acta de la reunión 6 se da por concluido, tras la realización de las tareas asignadas.	
4	Tareas <i>Sprint</i> 4 Se presentan las tareas del cuarto <i>sprint</i>	001
5	Otros asuntos No se tratan otros asuntos	
6	Próxima reunión Al completar todas las tareas del <i>sprint</i>	

Resumen de acuerdos

Número	Acuerdo	Plazo	Responsable
001	Realización de las tareas del <i>sprint</i> 4	03/07/2023	Marcos Marina

Reunión 9

REUNIÓN: Revisión de la memoria

Fecha: 29/06/2023	
Hora comienzo: 14:30	Hora finalización: 15:30
Lugar: Universidad San Jorge	
Elabora acta: Marcos Marina	
Convocados: Marcos Marina, Violeta Monasterio	

Orden del día / Acta

No.	Asunto	Acuerdo
1	Disculpas por ausencia No procede	
2	Aprobar última acta No procede	
3	Asuntos pendientes última acta No procede	
4	Correcciones en la memoria Se presentan las diferentes correcciones que se deben hacer en la memoria	001
5	Otros asuntos No se tratan otros asuntos	
6	Próxima reunión Al completar la memoria y realizar las correcciones	

Resumen de acuerdos

Número	Acuerdo	Plazo	Responsable
001	Realización de las correcciones	11/08/2023	Marcos Marina

Reunión 10

REUNIÓN: Fin *sprint* 4. Inicio *sprint* 5

Fecha: 03/07/2023	
Hora comienzo: 9:00	Hora finalización: 9:45
Lugar: Teams	
Elabora acta: Marcos Marina	
Convocados: Marcos Marina, Rafael del Hoyo	

Orden del día / Acta

No.	Asunto	Acuerdo
1	Disculpas por ausencia No procede	
2	Aprobar última acta No procede	
3	Asuntos pendientes última acta El acuerdo 001 del acta de la reunión 8 se da por concluido, tras la realización de las tareas asignadas.	
4	Tareas <i>Sprint</i> 5 Se presentan las tareas del quinto <i>sprint</i>	001
5	Otros asuntos No se tratan otros asuntos	
6	Próxima reunión Al completar la memoria y realizar las correcciones	

Resumen de acuerdos

Número	Acuerdo	Plazo	Responsable
001	Realización de las tareas del <i>sprint</i> 5	30/08/2023	Marcos Marina

Reunión 11

REUNIÓN: Últimos cambios de la versión preliminar

Fecha: 19/08/2023	
Hora comienzo: 15:30	Hora finalización: 17:30
Lugar: Teams	
Elabora acta: Marcos Marina	
Convocados: Marcos Marina, Violeta Monasterio	

Orden del día / Acta

No.	Asunto	Acuerdo
1	Disculpas por ausencia No procede	
2	Aprobar última acta No procede	
3	Asuntos pendientes última acta El acuerdo 001 del acta de la reunión 8 se da por concluido, tras la realización de las correcciones	
4	Correcciones en la memoria Últimas indicaciones para las correcciones en la memoria del PFG a cargo de Violeta Monasterio	001
5	Otros asuntos No se tratan otros asuntos	
6	Próxima reunión Al completar la memoria y realizar las correcciones	

Resumen de acuerdos

Número	Acuerdo	Plazo	Responsable
001	Correcciones y envío de la memoria a Alfonso Santolaria	20/08/2023	Marcos Marina

Reunión 12

REUNIÓN: Últimos cambios de la versión preliminar

Fecha: 24/08/2023	
Hora comienzo: 17:00	Hora finalización: 18:30
Lugar: Teams	
Elabora acta: Marcos Marina	
Convocados: Marcos Marina, Alfonso Santolaria	

Orden del día / Acta

No.	Asunto	Acuerdo
1	Disculpas por ausencia No procede	
2	Aprobar última acta No procede	
3	Asuntos pendientes última acta El acuerdo 001 del acta de la reunión anterior se da por concluido, tras el envío de la memoria	
4	Correcciones en la memoria Indicaciones para las correcciones en la memoria del PFG a cargo de Alfonso Santolaria	001
5	Otros asuntos No se tratan otros asuntos	
6	Próxima reunión Al completar la memoria y realizar las correcciones	

Resumen de acuerdos

Número	Acuerdo	Plazo	Responsable
001	Realización de las correcciones propuestas	01/09/2023	Marcos Marina

Reunión 13

REUNIÓN: Cierre del proyecto

Fecha: 01/09/2023	
Hora comienzo: 9:00	Hora finalización: 10:30
Lugar: Teams	
Elabora acta: Marcos Marina	
Convocados: Marcos Marina, Rafael del Hoyo, Violeta Monasterio, Alfonso Santolaria	

Orden del día / Acta

No.	Asunto	Acuerdo
1	Disculpas por ausencia No procede	
2	Aprobar última acta No procede	
3	Asuntos pendientes última acta El acuerdo 001 del acta de la reunión 10 se da por concluido, tras la realización de las tareas. El acuerdo 001 del acta de la reunión anterior se da por concluido, tras la realización de las correcciones	
4	Conclusiones Rafael expresa sus felicitaciones por el trabajo realizado por Marcos en el proyecto. Tanto Violeta como Alfonso compartieron la opinión de que se ha realizado un excelente trabajo. En respuesta, Marcos agradeció la colaboración y facilidad de trabajar con ellos en el proyecto. Además, se acordó que la memoria del proyecto será enviada según lo planeado.	001
5	Otros asuntos Preparación de la presentación	002
6	Próxima reunión Al completar la memoria y realizar las correcciones	

Resumen de acuerdos

Número	Acuerdo	Plazo	Responsable
001	Envío de la memoria	01/09/2023	Marcos Marina
002	Preparación de la presentación	12/09/2023	Marcos Marina

ANEXO III: Resultados ITAINNOVA

A continuación, se presenta una tabla que recopila los resultados obtenidos en el marco del proyecto ITAINNOVA. Esta recopilación de resultados constituye un recurso valioso para entender el rendimiento relativo de los modelos y sus adaptaciones a diferentes conjuntos de características. El resaltado en negrita proporciona una guía visual para identificar los logros más sobresalientes en cada métrica. Como se ha indicado con anterioridad, para esta evaluación se utilizaron los ensayos 1 y 2 del "FALL-UP dataset" para el entrenamiento, mientras que el ensayo 3 se utilizó para la obtención de resultados.

HAR Model	Features	F1	Accuracy	Precision
LSTM	all	0.718	0.848	0.738
LSTM	3d	0.787	0.888	0.795
LSTM	2d	0.692	0.864	0.702
LSTM	imus	0.702	0.916	0.718
LSTM	ankle	0.515	0.768	0.550
CNN	all	0.781	0.892	0.816
CNN	3d	0.760	0.836	0.797
CNN	2d	0.616	0.835	0.676
CNN	imus	0.655	0.833	0.685
CNN	ankle	0.664	0.795	0.699
Transformer	all	0.865	0.891	0.882
Transformer	3d	0.825	0.881	0.818
Transformer	2d	0.825	0.864	0.819
Transformer	imus	0.793	0.896	0.852
Transformer	ankle	0.651	0.785	0.679
Espinosa et al. [45]		0.7294	0.8226	0.7425
Martinez-Villasenor et al.[29]		0.712	0.951	0.718
Suarez et al. [46]		0.8367	0.893	0.8485

ANEXO IV: Material Complementario

Dentro del material complementario de este proyecto, se puede acceder al siguiente contenido:

- a. **README.txt:** Un archivo que proporciona una descripción detallada del contenido adicional.
- b. **REPOSITORIO:** Una carpeta que contiene la totalidad del contenido presente en el repositorio Git en la fecha de presentación del proyecto.
- c. **DIAGRAMAS:** En esta carpeta, se encuentran las Figuras 16, 17 y 18, que corresponden a los diagramas de Clases. Estas figuras se han incorporado al material complementario para hacer más sencilla su visualización.
- d. **VIDEOS:** En esta carpeta, se incluyen los vídeos de las figuras 23, 24, 34 y 35, que representan capturas extraídas de vídeos. Estos vídeos se han integrado aquí con el fin de permitir su reproducción y visualización.