

**EVALUACIÓN DE MODELOS DE VISIÓN POR COMPUTADOR EN VIDEO PARA LA
DETECCIÓN DE LA POSE HUMANA Y CAÍDAS**

Daniela Andrea Calvache Briceño

Trabajo Dirigido

Tutor
PhD(c). Oscar Julián Perdomo Charry



UNIVERSIDAD DEL ROSARIO
ESCUELA COLOMBIANA DE INGENIERÍA JULIO GARAVITO
PROGRAMA DE INGENIERÍA BIOMÉDICA
BOGOTÁ D.C
2020

AGRADECIMIENTOS

Con mucho cariño,
a mis seres queridos, a mis profesores y a mis amigos

*“Qué será,
será,
what ever will be, will be,
the future is not our to see,
what will be,
will be”*

- *Doris Day*

TABLA DE CONTENIDO

| | |
|--|-----------|
| INTRODUCCIÓN | 4 |
| Inteligencia artificial | 5 |
| Aprendizaje automático | 6 |
| Aprendizaje profundo o Deep learning | 6 |
| Aspectos éticos de IA | 8 |
| Detección de Pose o Pose Estimation | 8 |
| Antecedentes: Análisis de la marcha | 9 |
| Colocación de los marcadores y calibración de las cámaras | 10 |
| Detección automática de la pose con aprendizaje profundo | 10 |
| Detección de caídas | 12 |
| OBJETIVOS | 14 |
| General | 14 |
| Específicos | 14 |
| METODOLOGÍA | 15 |
| Pregunta de investigación | 15 |
| Hipótesis | 15 |
| Recursos | 15 |
| Cronograma de trabajo | 16 |
| Instrumentación (Materiales) | 17 |
| Métodos | 18 |
| Selección del modelo de Pose Detection a utilizar | 19 |
| Captura de videos en la casa de actividades de la vida diaria de la Universidad del Rosario | 19 |
| Rango articular de las articulaciones y cálculo de ángulos de las extremidades de manera offline | 20 |
| Procesamiento de los videos | 22 |
| Análisis estadísticos | 23 |
| Imprevistos | 24 |
| RESULTADOS | 25 |
| DISCUSIÓN | 32 |
| CONCLUSIONES | 34 |
| RECOMENDACIONES Y TRABAJOS FUTUROS | 35 |

| | |
|--------------------|-----------|
| REFERENCIAS | 36 |
| ANEXOS | 37 |
| ANEXO-A | 37 |
| ANEXO-B | 38 |
| ANEXO-C | 43 |

ÍNDICE DE FIGURAS

| | |
|---|----|
| Fig. 1. Arquitectura de las CNN, donde sus capas están interconectadas entre sí, para la detección y clasificación de bases de datos de imágenes etiquetadas..... | 9 |
| Fig. 2. Extracción de características y capas de clasificación de las CNN..... | 9 |
| Fig. 3. La colocación de los marcadores anatómicos de las extremidades inferiores según el método de Davis..... | 12 |
| Fig. 4. Los modelos de detección de pose..... | 13 |
| Fig. 5. Técnicas de segmentación del DL..... | 14 |
| Fig. 6. Sistema de detección de caída por medio de métodos de segmentación..... | 15 |
| Fig. 7. Sistema de caída implementado por la Escuela de Ingeniería Eléctrica y de Computación de Georgia, basado en la segmentación de la persona en el cuadro..... | 15 |
| Fig. 8. Cámara Sony HandyCam..... | 19 |
| Fig. 9. Módulo de la cámara Raspberry..... | 19 |
| Fig. 10. Xiaomi Redmi Note..... | 19 |
| Fig. 11. Metodología del estudio..... | 20 |
| Fig. 12. Anatomía del codo..... | 22 |
| Fig. 13. Anatomía de la rodilla..... | 22 |
| Fig. 14. Definición de etiquetas de los ángulos a calcular..... | 23 |
| Fig. 15. Implementación del sistema de detección de caídas..... | 24 |
| Fig. 16. Etiquetas de los ángulos de las extremidades calculados y los puntos equivalentes del modelo de detección de pose y caídas..... | 25 |
| Fig. 17. Muestra del análisis manual de los ángulos con Kinovea..... | 26 |
| Fig. 18. Resultados del video capturado formato h264..... | 27 |
| Fig. 19. Resultados del video capturado formato MTS..... | 27 |
| Fig. 20. Resultados del video capturado formato mp4..... | 28 |
| Fig. 21. Gráficos de dispersión de los ángulos calculados..... | 29 |
| Fig. 22. Cuadros donde el video capturado con el módulo de cámara de la Raspberry..... | 30 |
| Fig. 23. Cuadros de referencia de la prueba individual del formato MTS (Camara Sony HandyCam)..... | 31 |
| Fig. 24. Secuencia de la caída desde una silla formato MTS..... | 31 |
| Fig. 25. Prueba con varias personas de izquierda a derecha..... | 32 |
| Fig. 26. Secuencia de la caída de la persona 2 formato de video h264(Cámara módulo de la Raspberry)..... | 32 |

ÍNDICE DE TABLAS

| | |
|---|----|
| Tabla I. Enumeración de las actividades a realizar..... | 18 |
| Tabla II. Especificaciones de las cámaras utilizadas en las pruebas..... | 20 |
| Tabla III. Rango articular no patológico de las articulaciones de las extremidades a considerar..... | 23 |
| Tabla IV. Etiquetas de los ángulos de las extremidades calculados..... | 24 |
| Tabla V. Ángulos calculados para el cuadro de referencia de inicio..... | 28 |
| Tabla VI. Ángulos calculados para el cuadro de referencia del fin..... | 28 |
| Tabla VII. Error relativo de los ángulos teniendo en cuenta la prueba estadística no paramétrica de la mediana..... | 29 |
| Tabla VIII. Especificaciones de los videos de captura..... | 30 |
| Tabla IX. Tasa de predicción adecuada del modelo cuando ocurren obstaculizaciones en el ambiente de prueba..... | 30 |
| Tabla X. Especificaciones de los videos de captura antes de la normalización en tiempo..... | 30 |
| Tabla XI. Desempeño del modelo durante la caída del sujeto de prueba desde una silla..... | 31 |
| Tabla XII. Especificaciones de los videos de captura antes de la normalización en tiempo..... | 31 |
| Tabla XIII. Tasa de detección de la caída según el sistema condicional teniendo en cuenta cuadros multipersona..... | 32 |

1. INTRODUCCIÓN

En el presente documento se expone una evaluación sistemática del desempeño de diferentes dispositivos utilizados para emplear modelos de visión artificial aplicados a videos para la detección de pose y detección de caídas de una o varias personas. Lo anterior con el propósito de evaluar las variables a tener en cuenta a la hora de implementar dicho sistema en un espacio de uso cotidiano en tiempo real. El impacto esperado de este estudio comprende mejorar algunos aspectos de la vida diaria de las personas de la tercera edad las cuales son las más propensas a sufrir caídas de alto impacto, para evitar las consecuencias físicas, mentales, sociales y clínicas que estas provocan. De igual manera se espera apoyar a las personas dedicadas al cuidado de esta población vulnerable, y que se reduzca la carga física, emocional y ocupacional que requiere este oficio. Además, para conceder a los ancianos una vejez digna en sus hogares y evitar la necesidad de establecerse en un hogar geriátrico. Cabe mencionar que este proyecto de grado complementa un proyecto macro de investigación llamado “Herramienta automática para análisis de actividades cotidianas en un ambiente residencial a través de técnicas de inteligencia artificial”, cuyo consentimiento informado y aval ético, que respaldan lo realizado en este proyecto de grado, se encuentran en la sección de anexos.

En la primera sección del documento, se exponen conceptos teóricos que contextualizan el estudio. La segunda sección incluye la metodología a desarrollar, y seguidamente, en la tercera sección se exponen los resultados. Finalmente, se presenta la discusión de los resultados, junto con las conclusiones y recomendaciones para trabajos futuros.

1.1. Inteligencia artificial

Para comenzar, el concepto de Visión Artificial (VA) se da de manera paralela a la Inteligencia Artificial (IA) a mediados del siglo XX, con la necesidad de generar imágenes de mayor resolución para un procesamiento más eficiente de las mismas, con propósitos variados, desde el estudio de cuerpos celestes, hasta mejorar la capacidad de dictaminar diagnósticos de patologías. VA nace de manera similar a los modelos que la caracterizan, ya que analiza la funcionalidad del ojo humano, buscando imitar su funcionamiento, y cuyas diferentes acciones sean aplicadas mediante sensores o unidades de procesamiento. El desarrollo de procesadores y unidades de procesamiento gráfico (GPU) a finales de la década de los 80 y los 90, abrieron las puertas a la instauración de un sistema de procesamiento de información más robusto, con más capacidad y potencia, que permitiría realizar tareas cada vez más complejas en menos tiempo, y liberaría las oportunidades para el procesamiento visual de información de manera más eficiente que la inteligencia artificial había estado esperando.

El desarrollo de los algoritmos de inteligencia artificial ha generado un alto impacto de cambio y controversia desde sus primeros atisbos, debido a su capacidad de modificar en algún aspecto las actividades de la vida diaria de las personas y los sectores industriales, económicos y sociales. La evolución de la IA, se ha dado gracias también a la cantidad de recursos de código abierto, y bases de datos, que están disponibles y que se combinan con conocimientos de fotónicos e interfaces cerebro-máquina (BCI, por sus siglas en inglés)[12], para potenciar la innovación en el área. El diseño e implementación de las arquitecturas que propone son tan variadas y desafiantes, como los retos a los cuales pretenden dar solución. Entre las aplicaciones

más conocidas que utilizan los incipientes métodos de aprendizaje automático, se encuentra en primer puesto el traductor de Google. Este consta de un codificador y un decodificador, con capas neuronales ocultas entre ellas que permiten una traducción rápida y precisa que además va mejorando con su uso[10]. Otra aplicación es la conducción automática de los automóviles, que además cuentan con un sistema de detección de obstáculos, garantizando una desaceleración segura, y un asistente de parqueo como en algunas camionetas ya presentes en el mercado.

Específicamente en la medicina, se ha ganado bastantes acreedores debido a que ofrece herramientas confiables para el análisis automatizado de imágenes médicas en oncología, y el procesamiento de información en tiempo real haciendo uso de asistentes virtuales, entre otras cosas. El desarrollo de dispositivos vestibles (*wearable devices*) que se han desarrollado a la par, adquieren, almacenan y comparten la información, y aprovechan la interconectividad de los dispositivos inteligentes (IoT), y aplican conocimientos de *Big Data* para el procesamiento de grandes cantidades de información[12].

1.1.1. Aprendizaje automático

El aprendizaje automático (ML, por sus siglas en inglés, machine learning) está basado en el proceso de aprendizaje del ser humano, que es crítico en la temprana edad, en donde la experiencia se construye a través del ensayo y error. Lo que diferencia al ML de los demás algoritmos de programación, radica básicamente en que los modelos aprenden de los datos ciertos criterios y buscan encontrar patrones que puedan ser empleados en diferentes acciones. Por lo tanto, la resolución de los problemas no está determinada por un conjunto de instrucciones explícitamente programados[13].

El ML se divide actualmente en tres categorías según el set de datos que se utilizan para entrenar y validar los modelos[1]. La primera categoría, es el aprendizaje supervisado, que se caracteriza por métodos que utilizan un set de datos del cual se conoce la variable objetivo, se conoce la salida del sistema o dicho de otra manera, se tiene el vector objetivo de las predicciones del modelo. La segunda categoría, es el aprendizaje no supervisado, cuyos métodos usan sets de datos no etiquetados, cuya salida no se conoce, pero de igual manera se espera que el método encuentre las predicciones, mejorando los niveles de exactitud y disminuyendo su error en cada iteración o época[1]. La tercera categoría es el aprendizaje reforzado o por refuerzo, el cual combina un poco las dos clases explicadas antes, ya que cuando el modelo encuentra la respuesta correcta recibe una reafirmación de la respuesta apropiada, lo que es similar al aprendizaje supervisado. Más cuando es incorrecta, no recibe una corrección si no una evaluación del error, lo que es característico del aprendizaje no supervisado.

Dentro del aprendizaje automático, están las redes neuronales (NN, por sus siglas en inglés, neural networks) que están inspiradas en el funcionamiento biológico del cerebro cuyas zonas se especializan en tareas específicas y responden de una manera puntual dependiendo del estímulo que reciba el cerebro[7].

1.1.1.1. Aprendizaje profundo o Deep learning

Actualmente con los avances en las NN, la cantidad de procesamiento y el manejo de grandes volúmenes de información surgen las redes neuronales convolucionales (CNN), las cuales se componen de un conjunto de capas y parámetros que facilitan el entrenamiento de los modelos de clasificación y reconocimiento de bases de datos, evitando la necesidad de la extracción de las características de forma manual. Actualmente, existe gran cantidad de modelos desarrollados junto

con grandes bases de datos disponibles, que permiten desarrollar modelos con millones de parámetros y con desempeños sobresalientes.

En trabajos como el de Yann LeCun et. al., se exponen modelos de CNN, que han sido muy exitosos en la detección y clasificación de imágenes[16]. Como se observa en la Fig.1, los módulos o lo que serían las “neuronas” de la red se interconectan entre sí, comunicando a una neurona con una o varias de la capa anterior y posterior. Entre las desventajas del DL y más específicamente las CNN, es que el sobreajuste(*overfitting*) puede ser muy común si los datos no son lo suficientemente vastos[16]. El sobreajuste se da cuando el modelo es incapaz de generalizar ante las entradas que se le proponen.

Este tipo de redes se caracterizan por realizar un preprocesamiento al conjunto de datos con las primeras capas de la arquitectura de la red, que son una serie de capas convolucionales con filtros de imágenes conocidas como kernels convolucionales. En las CNN también se pueden incluir y combinar capas de pooling, como se muestra en la Fig. 2 las cuales normalizan y/o reducen la resolución de las imágenes de manera que la máquina aprende cuales son los mejores métodos para lograr predicciones más precisas y exactas. La exactitud de las predicciones se perfecciona con cada iteración gracias al método de Backpropagation, el cual varía el peso o importancia, de cada entrada de manera conveniente, que es aprendida por la máquina. Entre las CNN, está la red AlexNet, que se caracteriza por sus 8 capas; 5 convolucionales y 3 capas completamente conectadas(full-connected layer), la cual permitió alcanzar niveles de exactitud en la detección y clasificación de objetos que no se había visto antes, durante en una época en donde la comunidad había dejado de lado las CNN.

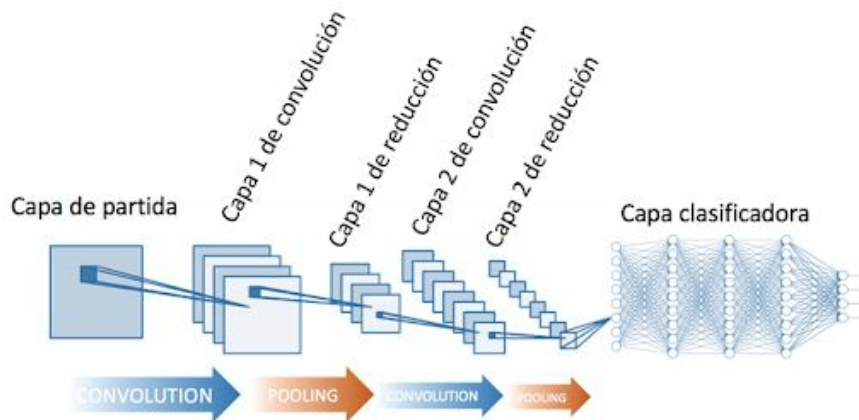


Fig. 1. Arquitectura de las CNN, donde sus capas están interconectadas entre sí, para la detección y clasificación de bases de datos de imágenes etiquetadas.¹

¹ Imagen extraída de Creative Commons, imágenes sin derechos de autor.

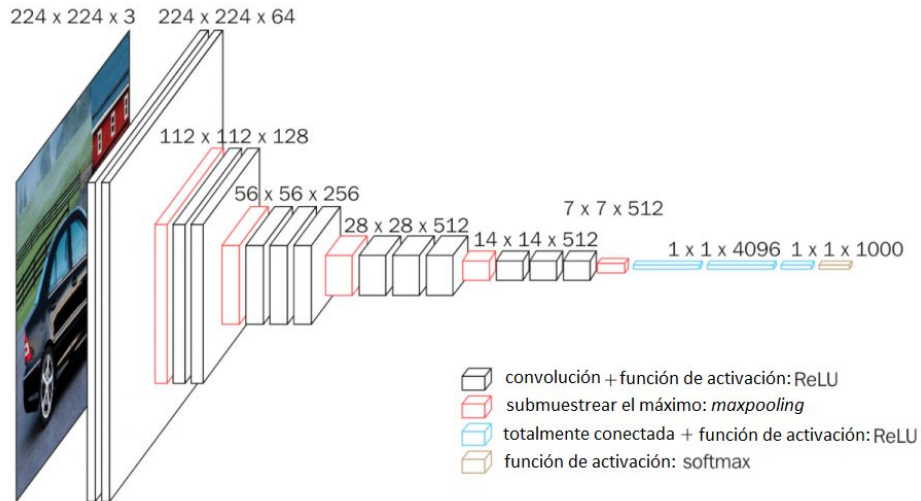


Fig. 2. Extracción de características: las capas convolucionales (negro) y con la salida representada en (ancho, alto, número de filtros), capas de max-pooling (rojo) que reducen la dimensionalidad de la salida de las convolucionales. Finalmente, la clasificación es desarrollada a través de capas densas(azules)[7].²

La aplicación de las redes neuronales comenzó por la identificación de caracteres en un sistema de lectura de códigos ZIP, y cheques, en Estados Unidos con Yann LeCun, a finales de los 90, después de que la comunidad perdiera el interés por las NN y lo recuperara a principios de los 2000. Gracias al desarrollo de sistemas más sofisticados y el auge del internet, se comenzó a innovar en un sistema de detección de objetos utilizando la reconstrucción de imágenes en estéreo[13]. En trabajos como el de Tang et. al. [2][11], se planteó la resolución de la detección de objetos como estructuras pictóricas, es decir donde un cuerpo se compone de sus partes. Los modelos planteados para la clasificación de la base de datos MNIST, de dígitos escritos a mano, y la clasificación de flores de IRIS, son actualmente reconocidos y usualmente trabajados para innovar en técnicas de aprendizaje automático. El trabajo en la detección de objetos llevó hacia la detección de pose, y a que se instaurase una comunidad dedicada a su desarrollo.

1.1.2. Aspectos éticos de IA

Siendo una tecnología que se expande con gran rapidez, y debido al manejo de datos sensibles que requiere para la implementación de los modelos, la IA ha sido un tema de gran controversia ética. En el Instituto de tecnología de Massachussets (MIT), la institución vanguardista en el área, se comienzan a plantear incógnitas sobre cómo ejercer un control sobre estas tecnologías emergentes para garantizar el bien común. Gómez, realizó una recopilación de las tecnologías relacionadas con la IA conocidas hasta el presente año, por medio de referencias representativas del área, y las clasifica en una escala de 0 a 9 dependiendo de su disponibilidad, discute sus fortalezas y debilidades, y sus beneficios a la sociedad según su grado de eticidad[12].

Por otro lado, el aprendizaje profundo demuestra una capacidad en crecimiento, y su aplicabilidad y versatilidad amenaza con ser indebidamente usada, como ha sucedido con los *deep fakes*. Cuya comunidad en la era de la interconectividad global explota a personalidades poderosas de la sociedad, las difaman y surgen con ellas nuevos contenidos junto con nuevos vocablos y

² Imagen extraída de Creative Commons, imágenes sin derechos de autor

anglicismos conocidos como *fake news* (noticias falsas). El transhumanismo o transformación al humanismo[19], se ha visto como un terreno desconocido y misterioso desde que el medio del entretenimiento ha explorado la ciencia ficción, y le ha ido mostrando al mundo las posibilidades de la IA mucho antes de que se desarrollaran. Con el avance de la tecnología que pareciera aún no llegar a su auge, Bostrom se plantea muy acertadamente, “...Son los programadores, los usuarios y finalmente los beneficiarios, quienes por distribución de responsabilidades y proximidad, deben asumir su inspección y evaluación (...), mediante una genuina y cualificada equidad en su aplicación –determinada por un panel de jueces humanos– escenario donde un algoritmo es aceptado en una aplicación transparente y rechaza negras intenciones”[19].

Por lo anterior, es importante reconocer que este estudio cuenta con el aval de la Universidad del Rosario para su desarrollo y el manejo de datos sensibles, el cual se anexa el final de este documento en la sección de anexos(Anexo-c).

1.2. Detección de Pose o Pose Estimation

La estimación de la pose humana, o HPE (por sus siglas en inglés, Human Pose Estimation [2]) es la identificación de una acción ejecutada por una o varias personas, la cual en una primera instancia, fue resuelta por los mismos métodos pictóricos de la detección de objetos y la posición de los segmentos entre las articulaciones del sujeto[2].

1.2.1. Antecedentes: Análisis de la marcha

De los primeros métodos implementados relacionados con la detección de pose, en trabajos como el de Leitch et. al., se emplea el uso de sensores inerciales y de plataformas de fuerza para el análisis de marcha[9]. El uso de los marcadores implica su exacta colocación en las articulaciones de las extremidades inferiores y puntos de referencia de la pelvis según el modelo de Davis de 1991, como se puede observar en Fig. 3[8]. Después de establecer los marcadores, se posicionan las cámaras convenientemente en un espacio cerrado, consiguiendo la detección de los sensores y su ubicación de manera digital, teniendo en cuenta un sistema de referencia.

La gran ventaja de este método es la confiabilidad de los resultados que arroja, además de que el protocolo está validado clínicamente, lo cual le confiere un nivel de certeza acogido por la comunidad, y lo categoriza como la prueba gold standard para la realización de esta prueba. La estimación de los puntos anatómicos de referencia permite la aproximación de la localización de las articulaciones, y con ello se realiza el análisis biomecánico del cuerpo mediante el análisis de marcha. Posteriormente, realizando la estimación de algunos parámetros, como; la longitud de zancada, el balance y la fuerza aplicada en intervalos de tiempo durante las fases de apoyo y levantamiento de talón, se puede realizar un diagnóstico aproximado del estado del sistema óseo y muscular de la persona. Los métodos tradicionales permiten digitalizar esta información y realizar un análisis manual con herramientas de software que se acoplan con el sistema de marcadores.

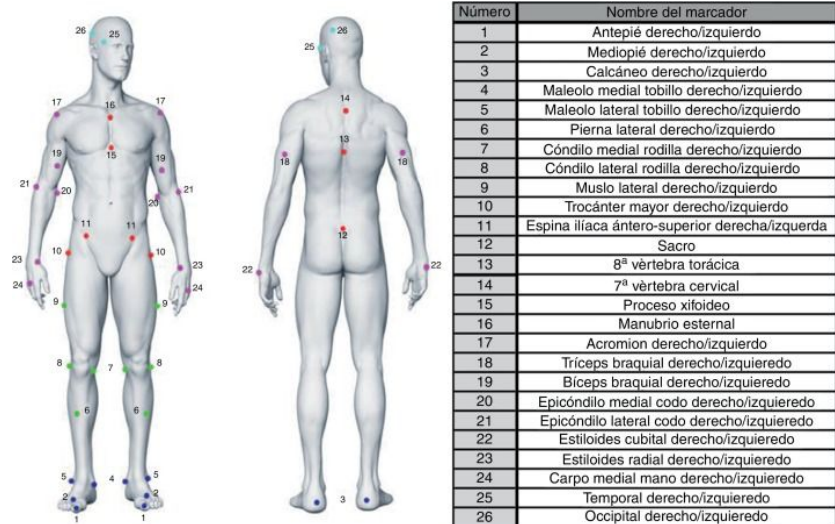


Fig. 3. La colocación de los marcadores anatómicos de las extremidades inferiores según el método de Davis están numerados en la figura así; 1, 2, 5, 7, 8, 10, 11 y 12.³

En la segunda parte del trabajo de Leitch, se realiza la detección de los tiempos de las fases de apoyo y levantamiento de talón, en ausencia de las plataformas de fuerza, teniendo en cuenta 4 modelos basados en la trayectoria, velocidad y aceleración de los tobillos[9]. Uno de los modelos utilizados es el algoritmo de la velocidad del pie (FVA, por sus siglas en inglés) que permiten la detección de las fases de la caminata. Sin embargo, la colocación de los marcadores es un método tedioso que incluye el riesgo de que se coloquen de manera incorrecta y la probabilidad de que se desajusten de su lugar apropiado durante las pruebas, además de que las medidas antropométricas necesarias varían con cada sujeto de prueba y la calibración de las cámaras debe realizarse en cada captura que se vaya a realizar. Por lo que a continuación, se explican con mayor detalle los métodos necesarios para la colocación y uso de los sensores en un laboratorio de marcha.

1.2.1.1. Colocación de los marcadores y calibración de las cámaras

La colocación de los marcadores se inicia realizando medidas antropométricas del sujeto de prueba, como; la amplitud de la pelvis, la longitud de las extremidades inferiores, el diámetro de la rodilla, diámetro del tobillo y la profundidad de la cadera, con ayuda de un metro y/o antropómetro, por lo que es necesario que el sujeto vista con las prendas adecuadas, para localizar los puntos anatómicos de manera más fácil. Además, se realiza la toma de sus datos corporales, como; edad, altura y peso. Los marcadores poseen un lado adhesivo, por lo que es necesario adecuar la piel del sujeto de prueba realizando el proceso de eliminación del vello, limpieza, y desinfección de la piel, y en algunos casos aplicar gel para disminuir la impedancia de la piel, dependiendo del tipo de sensores que se utilicen. Para ubicar los puntos del modelo de Davis, según la Fig. 3, se ubican; las crestas ilíacas, el sacro, los puntos más prominentes próximos a la rodilla, del fémur y la tibia, los tobillos, el empeine y el talón de ambas piernas.

Adicionalmente, es necesario calibrar el conjunto de cámaras presentes en el laboratorio de marcha en cada sesión que se fueran a realizar las pruebas. Estas se ubican de manera estratégica en el laboratorio y con un sistema de referencia definido.

³ Imagen extraída de Creative Commons, imágenes sin derechos de autor

1.2.2. Detección automática de la pose con aprendizaje profundo

Por otra parte, la HPE se puede realizar usando un video de uno o varios planos. Un video es una sucesión ordenada de imágenes en cada cuadro[2]. Las imágenes pueden ser de una vista(monoculares), de dos vistas(estéreo) o multivista. La clasificación de estas imágenes depende de la cantidad de cámaras estéreo utilizadas para capturarlas. Lo que caracteriza a las imágenes estéreo o multivista es que las cámaras se ubican a cierta distancia entre ellas para capturar las diferentes vistas de un evento en un mismo instante. Las ventajas de utilizar imágenes estéreo es la capacidad de combinar los resultados de la estimación de la pose de ambas lentes; izquierda y derecha, sin importar el modelo de ML a utilizar. De esta manera, se puede conseguir una realimentación para corregir las estimaciones obtenidas de ambas lentes [2].



Fig. 4. Los modelos de detección de pose permiten determinar puntos próximos a las articulaciones y partes del cuerpo, que son de interés de la persona o personas presentes en el cuadro para determinar su postura.⁴

Entre los métodos propuestos para la detección de pose de personas en video, Chunyu, en 2013, utilizó el modelo del aprendizaje no supervisado, 'K-means', para conseguir las K-mejores estimaciones de pose, incorporando métodos de segmentación semántica de imágenes, como se muestra en la Fig. 5b, y restricciones temporales[4]. De esta manera, se consiguen localizar de manera más precisa las articulaciones del cuerpo humano, teniendo en cuenta los modelos pictóricos de la detección de objetos que agrupan las articulaciones localizadas según la extremidad a la que pertenecen. Sin embargo, como se ve en trabajos como el de Toshev[3], lo más frecuente es encontrar que se utilizan CNN para la detección de pose, de manera que se identifica que cada una de las articulaciones encontradas hacen parte de un todo, y así se puede predecir la localización de una articulación que no esté visible o esté obstaculizada en el cuadro.

⁴ Imagen extraída y modificada de: <https://medium.com/@masherov14/pose-estimation-metrics-844c07ba0a78>

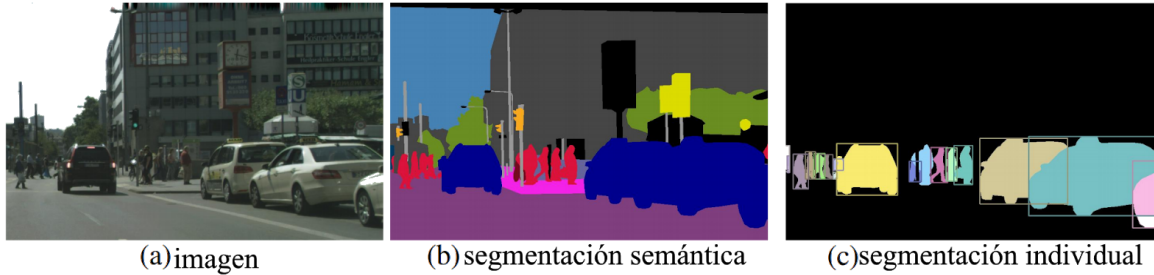


Fig. 5. Se muestran técnicas de segmentación del DL, aplicados a una imagen de una calle. b) Muestra la segmentación semántica, que asocia un categoría a cada pixel de la imagen, y c) la segmentación individual de cada objeto en el cuadro.⁵

La incorporación de las CNN para la detección de pose supuso grandes avances con respecto a la exactitud de la predicción. Se han implementado modelos tipo árbol, que evalúan la simetría y proporcionalidad del sujeto, que codifican la relación espacial que guardan las partes detectadas cercanas (las extremidades) y que siguen una cadena cinética[15]. Por otro lado, también están los modelos opuestos, que evalúan la oclusión de las partes, la simetría, y las relaciones espaciales de largo alcance.

1.3. Detección de caídas

Como se mencionó anteriormente, las caídas graves amenazan a las personas vulnerables de la tercera edad. Estas pueden perturbar su salud, causando graves fracturas que desembocan en cirugías invasivas, con el riesgo de infección más alto debido a la propia edad, estadías de hospitalización más largas que pueden generar desgaste intensivo muscular, escaras, infección e incluso la muerte. Por ende, la industria tecnológica ya produce sistemas de detección de caída para evitar estos sucesos y de paso aliviar de sistema de salud.

En el trabajo de Piñeros, realizado en la Fundación Cardio Infantil en conjunto con la Universidad del Rosario, se realizó un estudio de un sistema de detección en tiempo real de los cambios del movimiento de individuos sanos, utilizando sensores inteligentes de red inalámbrica con el propósito de prevenir las caídas. Junto con un prototipo de funcionamiento con una unidad de medición inercial, se detectan las variaciones que indican una inminente caída de manera previa. Teniendo en cuenta el estado del arte, en este trabajo se compararon diferentes estudios que involucraron el uso de sensores de presión adheridos a superficies frecuentemente utilizadas por los pacientes, como; la cama, o silla de una habitación. Además tuvieron en cuenta el estudio de Wolf, que hizo uso de un sensor inercial ubicado en el muslo de la persona[20]. Pero de nuevo, este tipo de equipo extracorpóreo incluye su colocación y calibración que requiere de cierto de tiempo de preparación y cuidados a personas no aptas para su uso.

De manera similar, y al igual que en la inteligencia artificial, los sistemas de detección de caída ya disponibles aprovechan los dispositivos móviles inteligentes vestibles (dispositivos e-health; relojes, bandas, manillas) para implementar alertas y comunicarse con teléfonos acoplados para enviar mensajes de emergencia en caso de una caída severa. Entre ellos está el Sense4care, que se expone en la Fig. 6, el cual es un dispositivo vestible, recargable con un acelerómetro triaxial que capta perturbaciones en cualquier eje, y cuenta con un botón de pánico.

⁵ Imagen modificada y extraída de:
<https://towardsdatascience.com/panoptic-segmentation-with-upsnet-12ecd871b2a3>

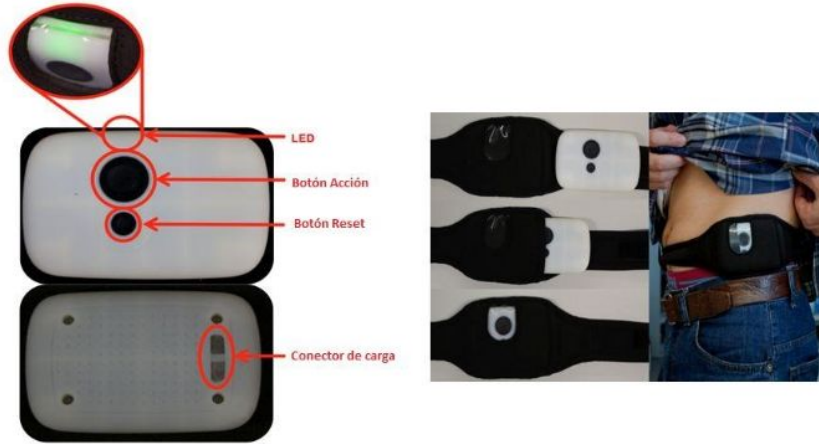


Fig. 6. Sistema de detección de caída Sense4Care, basado en un acelerómetro triaxial, con capacidad de interconexión con el teléfono para el envío de mensajes de emergencia.⁶

También existe el sistema de “Videovigilancia” de la Escuela de Ingeniería Eléctrica y de Computación de Georgia, basado en VA, que realizando segmentación a las imágenes, a partir de videos captados por cámaras posicionadas de manera conveniente en una habitación, captan los movimientos y calculan las variaciones anormales de velocidad del sujeto en el cuadro. El método de detección de caída se realiza por medio de un código en Matlab que segmenta a la persona en píxeles en escala de grises, como se observa en la Fig. 7, en donde el píxel blanco pertenece a la persona, el gris se genera con el movimiento, y el negro es el fondo, y se determina una probabilidad de caída de 0 a 1, en donde 0 es muy poco propenso a sufrir una caída, y 1, con gran probabilidad de caer basado en (1). Si se supera un umbral del 0.25 se dice que existe una caída.

$$P = \frac{N^{\circ} \text{ píxeles grises}}{N^{\circ} \text{ píxeles grises} + N^{\circ} \text{ de píxeles blancos}} \quad (1)$$



Fig. 7. Sistema de caída implementado por la Escuela de Ingeniería Eléctrica y de Computación de Georgia, basado en la segmentación de la persona en el cuadro. Píxeles negros son el fondo, los blancos son la persona, los grises se generan con el movimiento.⁷

⁶ Imagen extraída de: <http://www.alzfae.org/fundacion/461/como-elegir-un-detector-de-caidas>

⁷ Imagen extraída de: <http://www.alzfae.org/fundacion/461/como-elegir-un-detector-de-caidas>

2. OBJETIVOS

2.1. General

Realizar un análisis de detección de pose y detección de caídas, por medio de un sistema implementado con un modelo de visión por computador, utilizando 3 dispositivos diferentes para la captura de videos .

2.2. Específicos

1. Generar un conjunto de datos de capturas de video de varias personas realizando diferentes poses y acciones con diferentes dispositivos de captura.
2. Analizar las capturas y obtener la representación de las personas en un espacio durante el video.
3. Calcular los ángulos generados por el movimiento de las extremidades de la persona en cada cuadro del video y detectar patrones; antes durante y después de la caída.
4. Evaluar la capacidad del método propuesto en la detección automática de cada video, utilizando las tomas de video en diferentes ambientes y con diferentes dispositivos de captura.

3. METODOLOGÍA

3.1. Pregunta de investigación

¿Es posible implementar un sistema para la detección automática de pose y caídas de persona(s) de bajo costo adaptable a un ambiente de vida cotidiano?

¿El desempeño del modelo de detección de pose y de caídas propuesto, varía según el tipo de dispositivo empleado para realizar las capturas de video?

3.2. Hipótesis

Un sistema de detección de caída implementado con una cámara de video de teléfono inteligente gama media o con una tarjeta de desarrollo para realizar la captura de videos, puede ofrecer buena resolución y cuadros por segundo útiles en la detección de pose y caídas con un buen nivel de confianza.

3.3. Recursos

3.3.1. Framework TensorFlow. Google Collaboratory

TensorFlow es una plataforma de uso libre con librerías de código abierto en línea, desarrollado por el BrainTeam de Google junto con el equipo de inteligencia artificial(IA). Este framework, se utiliza principalmente en el estudio investigativo de modelos IA y redes neuronales de aprendizaje profundo[6].

3.3.2. Base de datos: MPII Human Pose Dataset

Esta base de datos crea un punto de referencia para la estimación de pose, y contiene aproximadamente 25,000 imágenes con anotaciones de las articulaciones del cuerpo de al menos 40,000 personas y un total de 410 actividades humanas ejecutadas. Las imágenes allí contenidas se seleccionaron teniendo en cuenta una clasificación previamente establecida de las actividades diarias del ser humano, clasificadas en 410 categorías. Las imágenes fueron extraídas de Youtube y se adicionaron anotaciones que especifican la obstaculización de partes del cuerpo, y las orientaciones en 3 dimensiones de la cabeza y el torso[5].

3.3.3. Modelos de estimación de pose: Open Pose y Pose Detection

El modelo de Open Pose está basado en campos de afinidad de las partes (PAF's), con los que se le asignan a un sujeto detectado. Se recibe como entrada una imagen o video directamente en la CNN que compone el modelo, ésta predice mapas de confianza para la detección de las partes del cuerpo, los campos de afinidad de las partes para la asociación de las partes a un sujeto, y en el paso de análisis, se realiza un conjunto de coincidencias de dos partes para asociar candidatos a partes del cuerpo detectadas. Este modelo fue diseñado por el equipo de investigación

de la Universidad Carnegie Mellon, en Pennsylvania. Utilizan las bases de datos MPII y COCO para entrenar y validar su modelo[15].

Por el otro lado, algunas alternativas como el modelo Pose Detection ofrece 4 configuraciones que dependen del procesamiento disponible; *CMU*, *mobilenet_thin*, *mobilenet_v2_large*, *mobilenet_v2_small*, que son parte de los requisitos de instalación para usar el modelo y permiten ejecutarlo en diferentes dispositivos y plataformas, y genera los puntos como se muestran en la Fig.5[17].

3.4. Cronograma de trabajo

En la tabla I, se listan las actividades enumeradas, realizadas para cumplir con los objetivos planteados en la sección 2, que se representa en el cronograma de trabajo el cual se anexa al final de este documento.

Tabla I. Enumeración de las actividades a realizar

| Número | Actividad | Fecha de inicio | Fecha de finalización |
|--------|--|-----------------|-----------------------|
| 1 | Revisión de literatura | 21/01/2020 | 8/02/2020 |
| 2 | Planteamiento de los objetivos | 5/02/2020 | 19/02/2020 |
| 3 | Primera entrega del documento | 24/02/2020 | 24/02/2020 |
| 4 | Selección del modelos de pose detection | 9/03/2020 | 11/03/2020 |
| 5 | Toma de las pruebas | 12/03/2020 | 30/03/2020 |
| 6 | Segunda entrega del documento | 30/03/2020 | 24/03/2020 |
| 7 | Procesamiento de los datos con Open Pose | 11/04/2020 | 14/04/2020 |
| 8 | Análisis de los resultados | 24/04/2020 | 21/01/2020 |
| 9 | Documento de trabajo de grado | 2/05/2020 | 24/04/2020 |
| 10 | Tercera entrega del documento | 2/05/2020 | 2/05/2020 |
| 11 | Entrega del documento al comité | 10/05/2020 | 10/05/2020 |
| 12 | Sustentación | 22/05/2020 | 22/05/2020 |

3.5. Instrumentación (Materiales)

La toma de las pruebas se realiza en el laboratorio de terapia ocupacional de la Universidad del Rosario llamado “Casa de actividades de la Vida diaria”. Se hace uso de las siguientes 3 cámaras para la captura de videos:

1. Módulo de cámara Raspberry
2. Cámara Sony ref. handycam
3. Cámara de smartphone Xiaomi Redmi Note 7.



Fig. 8. Cámara Sony HandyCam



Fig. 9. Módulo de la cámara Raspberry⁸



⁸Imagen extraída de:
<https://hardzone.es/2018/08/25/como-configurar-modulo-camara-raspberry-pi-3/>

Fig. 10. Xiaomi Redmi Note⁹

Las especificaciones de las cámaras empleadas en este trabajo se exponen a continuación en la tabla II.

Tabla II. Especificaciones de las cámaras utilizadas en las pruebas

| Tipo de cámara | Resolución fija | Modos de video | Resolución de imagen | Sensor | Tamaño del pixel | Cuadros por segundo | Formato de video | Precio (USD) |
|------------------------------|-----------------|--------------------|----------------------|---|-------------------|---------------------|------------------|------------------|
| Módulo de cámara Raspberry | 8 MP | 1080p30 | 3280 x 2464 pixels | Sony IMX219 | 1.12 μm x 1.12 μm | Hasta 120fps | h.264 | \$54.14(RB)+\$25 |
| Cámara Sony ref. Handycam 4K | 8,29 MP aprox. | 1080p30 | 3840 x 2160 pixeles | Sensor CMOS Exmor R® retroiluminado de tipo 1/2,5 (7,20 mm) | - | 100fps | MTS | \$869.85 |
| Xiaomi Redmi Note 7. | 5 MP | 1080p@30/60/120fps | 4000 x 3000 pixels | Exmor-RS CMOS Sensor | 1.12 μm x 1.12 μm | 30 60 120 fps | mp4 | \$231.45 |

3.6. Métodos

En la Fig. 11 se expone el proceso llevado a cabo para el desarrollo de este estudio. En las secciones siguientes se detallan las fases desarrolladas para el cumplimiento de esta investigación.

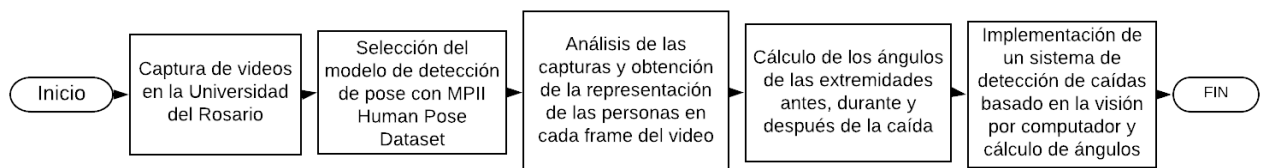


Fig. 11. Metodología del estudio

⁹Imagen extraída de:

<https://www.tiendasjumbo.co/Sistema/404?ProductLinkNotFound=celular-xiaomi-redmi-note-7-64gb-ds-4g-negro-20035138>

¹⁰ Información obtenida en <https://www.raspberrypi.org/documentation/hardware/camera/>,

<https://www.sony.es/electronics/handycam-camcorders/fdr-ax53/specifications>

https://www.gsmarena.com/xiaomi_redmi_note_7s-9707.php

3.6.1. Selección del modelo de Pose Detection a utilizar

Con el propósito de escoger un modelo de detección de pose para la implementación del sistema de detección de caídas, se utilizan imágenes de la base de datos MPII Human Pose Dataset, de una y varias personas para determinar el modelo con mejor desempeño.

La prueba se realiza para el modelo de Open Pose y Pose Estimation. En esta prueba se utiliza el modelo pequeño(*small*) de Pose Estimation y el modelo *Body_25*. Se observan de manera empírica los resultados de cada modelo. Se escogen 20 imágenes al azar de la base de datos Base de datos: MPII Human Pose Dataset, de actividades cotidianas que impliquen poses no naturales similares a una caída(bailar,hacer deporte: patinar, esquiar, nadar).

Se observan las salidas de los modelos, y se escoge el que menor tasa de error haya obtenido. La tasa de detección se define por indicadores cuantitativos y cualitativos de la detección, definiendo si el modelo detecta la pose adecuadamente, es decir si define los eslabones del cuerpo de manera apropiada, y si localiza los puntos de interés en las zonas de manera exacta.

3.6.2. Captura de videos en la casa de actividades de la vida diaria de la Universidad del Rosario

Se procede a realizar las capturas de video en la casa de la vida diaria de la Universidad del Rosario. Este espacio es destinado a realizar pruebas de prototipos e investigaciones con pacientes de terapia ocupacional y fisioterapia del departamento de la Universidad del Rosario.

Se ubican las cámaras; Sony, el módulo de cámara de Raspberry Pi 3 y la cámara del smartphone Xiaomi Redmi Note 7, en un mismo lugar estratégico de la casa a una altura aproximada de 1,70m y una distancia de 30 cm entre sí. Se realizan capturas de videos individuales, y con un grupo de 3 personas de aproximadamente 30 segundos.

Las tomas de pruebas se organizan así:

a) Pruebas individuales:

- i) El sujeto que simula una caída y se vuelve a poner de pie.
- ii) El sujeto utiliza una silla de ruedas y se desplaza por el lugar.
- iii) El sujeto simula se sienta en una silla y simula una caída.

b) Prueba con varias personas:

- i) Una de los sujetos simula una caída y mientras que los otros dos caminan por el lugar e interactúan con los objetos.

Los sujetos de prueba deben ser personas sin limitaciones físicas, sin un rango de edad específico y que no se vean afectados por caídas de leve impacto. Las prendas de vestir no son un factor ya que se desea evaluar la detección de pose y caída considerando un ambiente cotidiano simulado.

La primera prueba individual describe el sistema de detección de caídas con el modelo de detección de pose. Se realizarán pruebas comparativas que permitan evaluar el desempeño de las 3 cámaras especificadas en la tabla II. Las siguientes pruebas que incluyen más variables como la cantidad de personas, y la obstaculización de las articulaciones con los objetos, comparan el

desempeño del modelo utilizando las cámaras de mayor resolución fija; la Sony y el módulo de la Raspberry.

3.6.3. Rango articular de las articulaciones y cálculo de ángulos de las extremidades de manera *offline*

El método de detección de caída estará basado en el cálculo de los ángulos de las extremidades como se detalla posteriormente. Por lo que es necesario definir algunos conceptos previos relacionados con los codos y las rodillas.

Primero, el codo es una articulación en bisagra o pivote conformada por el cúbito, el radio y el húmero. Tanto el codo como la rodilla no es una sola articulación, si no 3, por las uniones de los extremos proximales y distales de los huesos que la conforman, como se muestra en la Fig. 12. Sin embargo, se dice que es una sola articulación por que funcionan como una unidad ya que las uniones están envueltas por una sola cápsula articular. Al igual que el codo, la rodilla se conforma por 3 huesos; el extremo distal del fémur, y los extremos distales de la tibia y el peroné, como se expone en la Fig. 13[18].

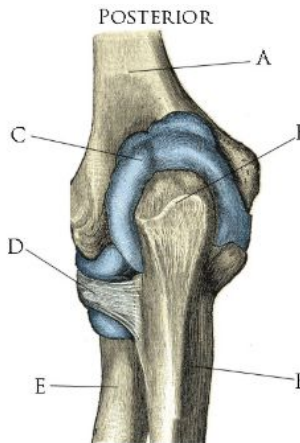
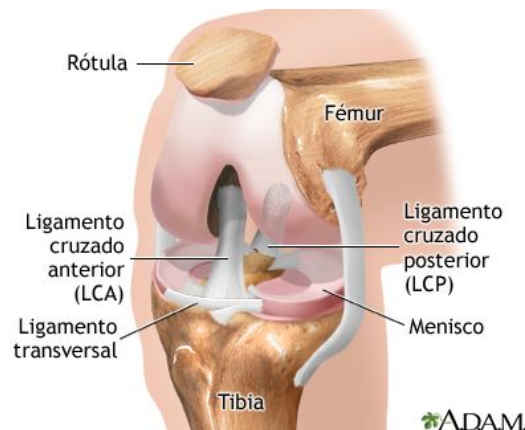


Fig. 12. Anatomía del codo, se señalan; A: Húmero, B: Olécranon, C: Bursa, D: Ligamento anular, E: Radio, F: Cúbito o ulna¹¹.



¹¹ Imagen extraída de: [18]

Fig. 13. Anatomía de la rodilla¹².

Tabla III. Rango articular no patológico de las articulaciones de las extremidades a considerar

| Articulación | Rango articular de flexo extensión |
|--------------|------------------------------------|
| Codo | 20°-180° |
| Rodilla | 30°-130° |

El cálculo de ángulos se realiza teniendo en cuenta la ley de cosenos aplicado al triángulo generado por la flexo-extensión de extremidades superiores o inferiores, que se aplica por cada cuadro del video. Los rangos articulares principales a calcular se exponen en la tabla III. Después, se define una función que recibe como parámetros los tres puntos que generan el ángulo. El ángulo se calcula mediante ley de cosenos con (2), en donde el vértice A corresponde al punto cercano a la articulación más proximal, el vértice C corresponde al punto cercano a la articulación más distal, y de esta manera el vértice B corresponde al punto entre A y C, en donde se genera el ángulo a medir, que en el caso de la Fig. 14, es el ángulo etiquetado como β . En el caso en que el modelo no detecte alguno de los puntos necesarios para el cálculo de los ángulos, se agrega una excepción y se define el ángulo como 0.0°.

$$\begin{aligned}
 a &= \text{dist}_{B,C} = \overline{BC} \\
 b &= \text{dist}_{A,C} = \overline{AC} \\
 c &= \text{dist}_{A,B} = \overline{AB} \\
 b^2 &= a^2 + c^2 - 2ac * \cos(\beta)
 \end{aligned}
 \tag{2}$$

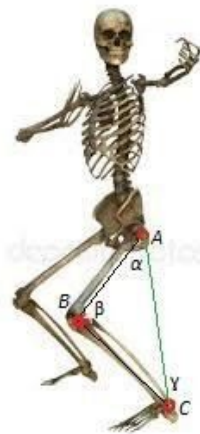


Fig. 14. Definición de etiquetas de los ángulos a calcular.¹³

¹² Imagen extraída de: https://medlineplus.gov/spanish/ency/esp_presentations/100117_1.htm

¹³ Imagen extraída y editada de: <https://sp.depositphotos.com/stock-photos/esqueleto-sentado.html>

3.6.4. Procesamiento de los videos

La metodología del procesamiento de los datos realizado y el sistema de detección de caída, dado el análisis con el modelo de estimación de pose se expone en la Fig. 15.

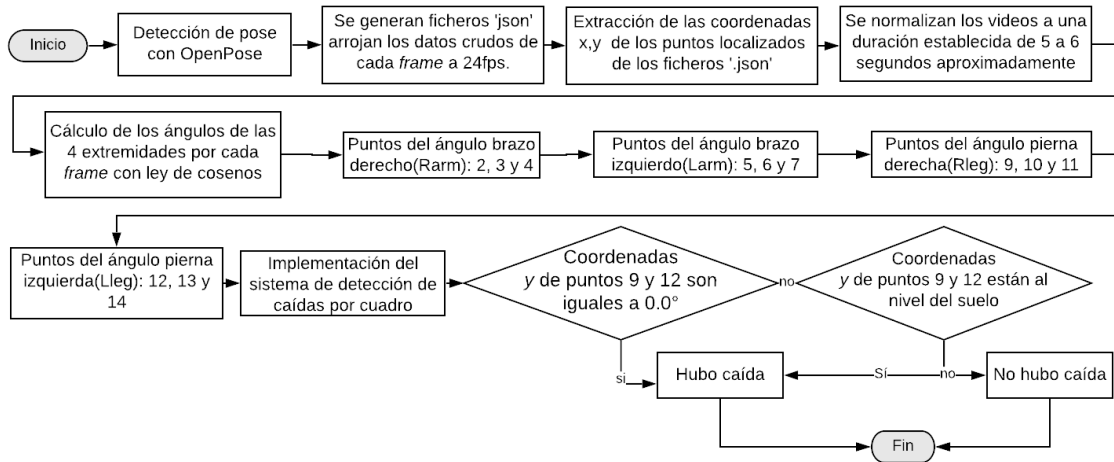


Fig. 15 Implementación del sistema de detección de caídas

Los valores de los ángulos calculados por cuadro, se denotan como se muestra en la tabla IV y en la Fig. 16, y se registran y almacenan en un *dataframe* con su respectiva etiqueta, para su posterior evaluación a partir del sistema de detección de caídas implementado. El sistema de detección de caídas se determinó basado en el desempeño del modelo de detección de pose utilizado, con un condicional teniendo en cuenta los ángulos calculados como se muestra en la Fig.15.

El inicio de la caída tomado como real, se define como el momento en que la rodilla de la persona toca el suelo, y el final de la caída es cuando la persona apoya las dos plantas de los pies en el suelo. Las coordenadas 'y' de los puntos 11 y 14 que equivalen a los puntos más proximales de los tobillos no varían. Adicionalmente se despliega una alerta en el video que indica la caída del sujeto, desde que la condición es verdadera hasta que deja de serlo, es decir hasta que el modelo es capaz de detectar la pose del sujeto de nuevo.

Tabla IV. Etiquetas de los ángulos de las extremidades calculados

| Anotación | Significado |
|-----------|------------------|
| Rarm | Brazo derecho |
| Larm | Brazo izquierdo |
| Rleg | Pierna derecha |
| Lleg | Pierna izquierda |



Fig. 16 Etiquetas de los ángulos de las extremidades calculados y los puntos equivalentes del modelo de detección de pose y caídas.

3.6.5. Análisis estadísticos

Para la evaluación de los datos calculados se hallan las siguientes medidas descriptivas de los resultados por formato; h264, MTS y mp4, y por cuadro.

- Media
- Varianza
- Desviación estándar
- Error absoluto y relativo

Se realiza la prueba estadística no paramétrica de la mediana, basado en varias muestras relacionadas, utilizada cuando no se conocen los datos exactos

3.6.6. Imprevistos

Dada la coyuntura de la pandemia del virus COVID-19, se vio limitado el uso del laboratorio de marcha dispuesto para la evaluación del estudio por las medidas gubernamentales relacionadas con el aislamiento preventivo. Por lo que se evaluaron algunas alternativas para el análisis manual de las capturas de video, entre ellas:

1. Kinovea
2. Skillspector
3. Tracker

En las Fig. 17, se presenta como se realiza la medición manual de los ángulos de un frame con el software Kinovea.



Fig. 17 Muestra del análisis manual de los ángulos con Kinovea

4. RESULTADOS

4.1. Prueba individual

4.1.1. El sujeto simula una caída y se vuelve a poner de pie

Los videos resultantes con la detección de pose, se normalizan a 4,29 segundos, lo que equivale a 103 cuadros, ya que se escogen los cuadros de inicio y fin de manera que sean los mismos en los 3 tipos de formatos a comparar. Este paso es necesario para obtener los mismos eventos a la tasa de muestreo equivalente. A continuación, se muestran una secuencia cuadros en las Fig. 18, 19 y 20 de los videos resultantes a partir de los videos capturados, de los formatos; h264, MTS y mp4, respectivamente en cada caso, en las que se muestran; el cuadro de inicio del video, el cuadro final y el primer cuadro en donde se registra caída del sujeto de prueba. En cada uno de los cuadros, se muestran los resultados de los ángulos calculados para Rarm, Larm, Rleg, Lleg. En las tablas V, y VI, se muestran los ángulos correspondientes a cada cuadro de referencia.

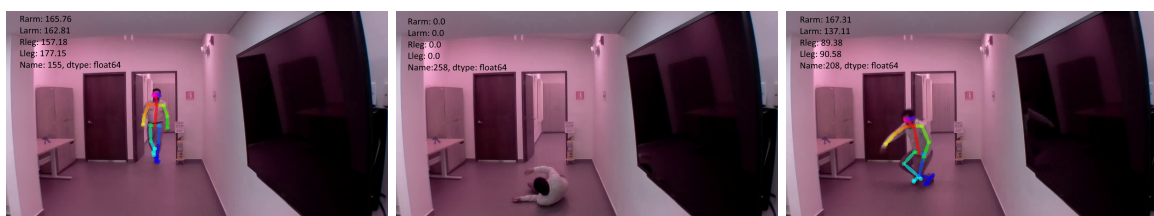


Fig. 18. Resultados del video capturado formato h264(módulo de la cámara Raspberry), a) cuadro de inicio, b) final del video, y c) primer cuadro de detección de caída.



Fig. 19. Resultados del video capturado formato MTS(cámara Sony), a) cuadro de inicio, b) final del video, y c) primer cuadro de detección de caída.



Fig. 20. Resultados del video capturado formato mp4(cámara del teléfono inteligente Xiaomi), a) cuadro de inicio, b) final del video, y c) primer cuadro de detección de caída.

Tabla V. Ángulos calculados para el cuadro de referencia de inicio

| Formato del video | Frame | Rarm | Larm | Rleg | Lleg |
|-------------------|-------|---------|---------|---------|---------|
| h264 | 1 | 165.76° | 162.81° | 157.18° | 179.14° |
| MTS | 1 | 176.61° | 174.62° | 143.34° | 170.19° |
| mp4 | 1 | 148.32° | 170.76° | 165.42° | 178.75° |

Tabla VI. Ángulos calculados para el cuadro de referencia del fin

| Formato del video | Frame | Rarm | Larm | Rleg | Lleg |
|-------------------|-------|--------|--------|--------|-------|
| h264 | 103 | NaN | NaN | NaN | NaN |
| MTS | 103 | 116.78 | 154.95 | 145.59 | 156.2 |
| mp4 | 103 | 143.43 | 179.33 | 15.16 | 81.11 |

En la Fig. 21 se expone el patrón de los ángulos calculados por extremidad dependiendo del formato de video.

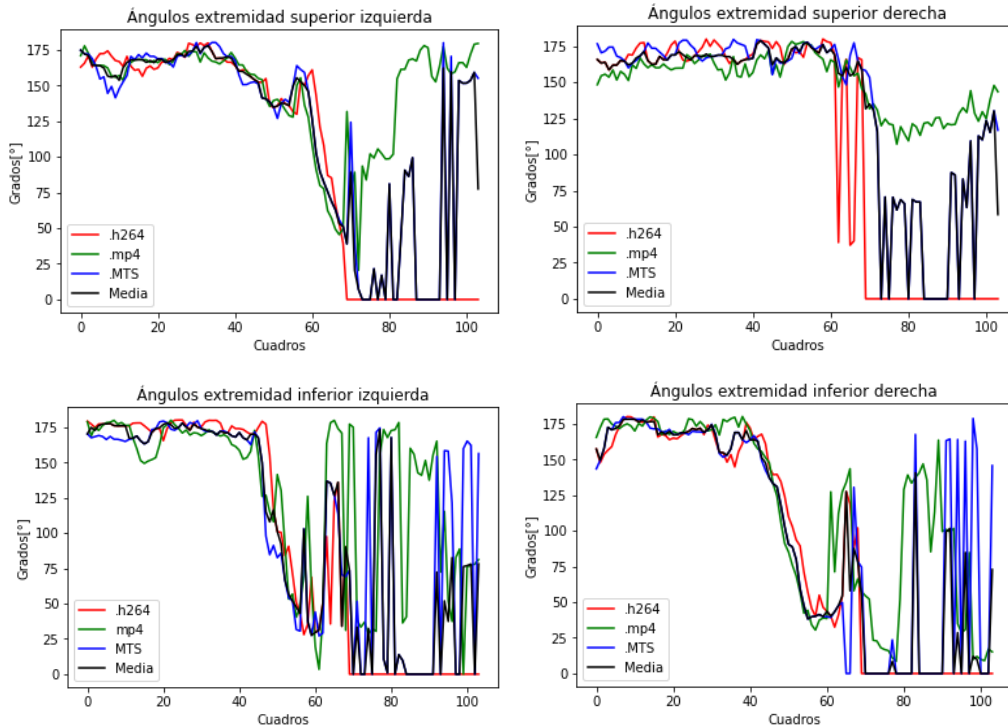


Fig. 21. Gráficos de dispersión de los ángulos calculados con respecto a la mediana por cuadros y por extremidad.

En la tabla VII, se indican los resultados estadísticos de dispersión por extremidad y en promedio, por cada formato del video que indican los valores normales de los cálculos realizados de los ángulos.

Tabla VII. Error relativo de los ángulos teniendo en cuenta la prueba estadística no paramétrica de la mediana

| Error Relativo(%) | Rarm | Larm | Rleg | Lleg | Media |
|-------------------|-------|-------|-------|-------|-------|
| h264 | 34.69 | 32.08 | 32.33 | 32.21 | 32.83 |
| MTS | 3.31 | 12.56 | 15.80 | 8.49 | 10.04 |
| mp4 | 38.00 | 44.65 | 48.14 | 40.71 | 42.87 |

4.1.2. El sujeto utiliza una silla de ruedas y se desplaza por el lugar

En la segunda prueba individual se pone a prueba la predicción del modelo cuando las articulaciones están obstaculizadas por la silla de ruedas que utiliza el sujeto de prueba. En la tabla VIII y se describe el desempeño del modelo cuantitativamente, donde de manera manual se encuentran los cuadros en donde se observa que el modelo falla realizando la predicción. Se comparan las capturas realizadas por el módulo de la cámara de la Raspberry y la cámara Sony HandyCam. En la Fig. 22 se exponen momentos en donde el modelo predice la localización de las articulaciones obstaculizadas por la vista de la cámara o de la silla de ruedas.

Tabla VIII. Especificaciones de los videos de captura.

| Formato del video | Duración del video (mm:ss:ms) | Cuadros | Cuadros detectados | Tasa de detección de pose | Caídas durante el video | Caídas detectadas |
|-------------------|-------------------------------|---------|--------------------|---------------------------|-------------------------|-------------------|
| MTS | 00:49 | 1468 | 1468 | 100% | 0 | 0 |
| h264 | 00:35 | 899 | 899 | 100% | 0 | 0 |



Fig. 22. Cuadros donde el video capturado con el módulo de cámara de la Raspberry (h264) predijo a) la posición de las extremidades inferiores obstaculizadas por la silla, b) los puntos cercanos a los tobillos cuando están obstaculizados por la vista de la cámara, c) predijo la posición de las extremidades obstaculizadas por la silla de ruedas. También se puede ver que el televisor refleja la imagen y distorsiona los datos verdaderos.

En la tabla IX, se expone la tasa de predicción adecuada del modelo cuando se presentan las obstaculizaciones de las extremidades de la Fig.27.

Tabla IX. Tasa de predicción adecuada del modelo cuando ocurren obstaculizaciones en el ambiente de prueba.

| Formato de video | Total de cuadros | Total de cuadros fallidos | Tasa de predicción correcta |
|------------------|------------------|---------------------------|-----------------------------|
| MTS | 1468 | 226 | 25,13% |
| h264 | 899 | 181 | 20,13% |

4.1.3. El sujeto simula una caída sentado en una silla

En esta prueba el sujeto se encuentra sentado en una silla, y simula caer de ella. En la tabla X, se muestran las especificaciones del video según los dos formatos a comparar; h264 y MTS.

Tabla X. Especificaciones de los videos de captura antes de la normalización en tiempo.

| Formato del video | Duración del video (mm:ss:ms) | Caídas durante el video | Caídas detectadas |
|-------------------|-------------------------------|-------------------------|-------------------|
| MTS | 00:32 | 1 | 3 |
| h264 | 00:00:35 | 1 | 0 |

En las Fig. 23 y 24, se exponen los cuadros de inicio y fin, y algunos cuadros de la secuencia de la caída. Por último, en la tabla XI se muestra el desempeño del modelo durante la caída y la detección de la misma.

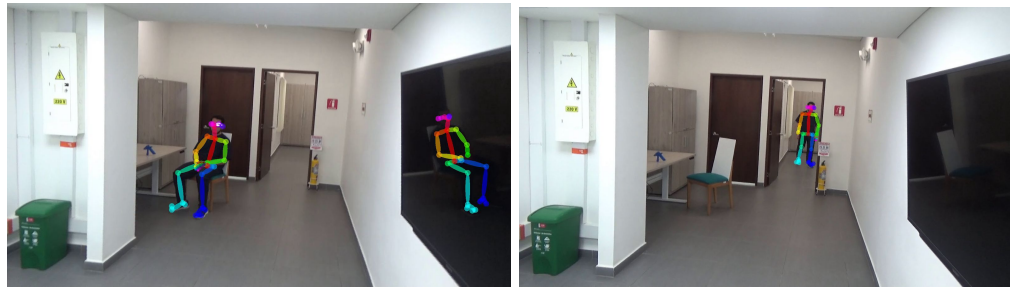


Fig. 23. Cuadros de a) inicio (Cuadro 0) y fin de la prueba (Cuadro 989) formato MTS (Camara Sony HandyCam)



Fig. 24. Secuencia de la caída desde una silla. Cuadros: a) 38, B)46, c)113.

Tabla XI. Desempeño del modelo durante la caída del sujeto de prueba desde una silla.

| Formato de video | Total de cuadros | Total de cuadros fallidos | Tasa de detección de pose |
|------------------|------------------|---------------------------|---------------------------|
| MTS | 989 | 0 | 100% |
| h264 | 898 | 5 | 99,44% |

4.2. Pruebas con varias personas:

4.2.1. Uno de los sujetos simula una caída, mientras que los otros dos caminan por el lugar.

En la Tabla XII, se exponen las especificaciones de los videos de las capturas con varias personas (3 personas), realizadas con la cámara Sony y el módulo de la Raspberry. En la Fig. 25 y en la Fig. 26, se enumeran las personas del video, y se muestra una secuencia de cuadros de la caída de la persona 2, respectivamente. Por último, en la tabla XIII, se expone la tasa de detección de caída obtenida.

Tabla XII. Especificaciones de los videos de captura antes de la normalización en tiempo.

| Formato del video | Duración del video (mm:ss:ms) | Cuadros | Caídas durante el video | Caídas detectadas |
|-------------------|-------------------------------|---------|-------------------------|-------------------|
| | | | | |

| | | | | |
|------|-------|------|-------------------|-------------------------------------|
| MTS | 00:49 | 1363 | 1 de la persona 2 | 1 (P1*) 14 cuadros (P2*) |
| h264 | 00:35 | 899 | 1 de la persona 2 | 2 cuadros (P1*) 75 cuadros (P2*) |



Fig. 25 Prueba con varias personas de izquierda a derecha; persona 1, persona 2 y persona 3.

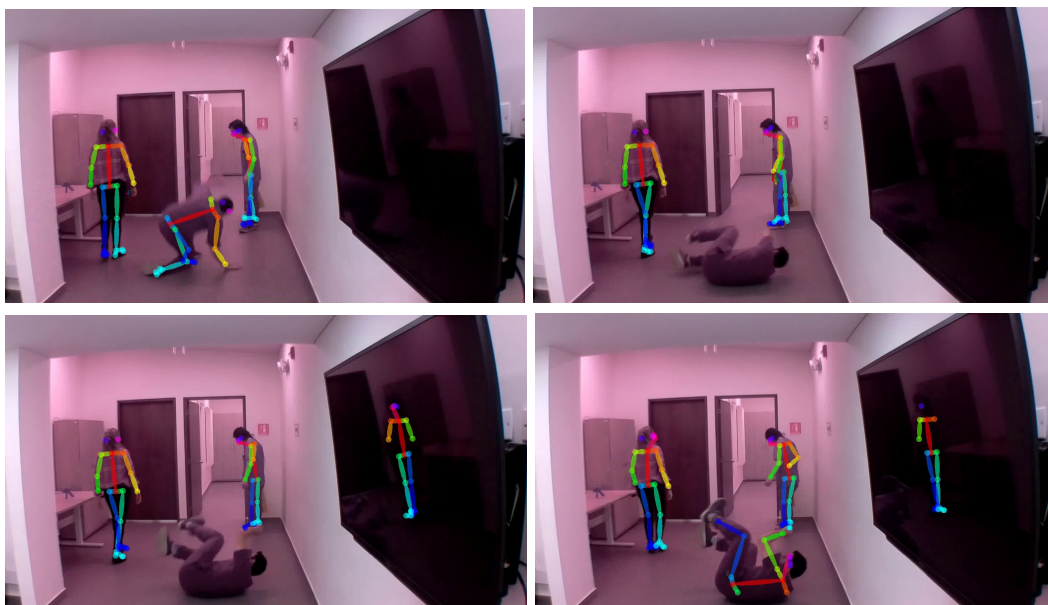


Fig.26. Secuencia de la caída de la persona 2 formato de video *h264*(Cámara módulo de la Raspberry). En el cuadro (a) se cumple la condición de caída 1(las coordenadas de los puntos 9 y 12 que equivalen a las rodillas a nivel del piso). Se detecta una caída.

Tabla XIII. Tasa de detección de la caída según el sistema condicional teniendo en cuenta cuadros multipersona.

| Formato de video | Total de cuadros | Cuadros en donde el modelo falla por persona | Total de cuadros fallidos | Tasa de detección de todas las personas en el video |
|------------------|------------------|--|---------------------------|---|
| MTS | 1362 | 12(P2*) + 248(P3*) | 260 | 80,91% |

| | | | | |
|------|-----|----------|----|--------|
| h264 | 898 | 56 (P3*) | 56 | 93,76% |
|------|-----|----------|----|--------|

**P1: persona 1, **P2: persona 2, **P3: persona 3

4.3. Análisis con Kinovea

En la Fig. 27, se exponen algunos cuadros de la prueba individual evaluada con el *software* Kinovea. El análisis manual de los ángulos de las extremidades se realiza con la opción de 'ángulo' del *software*, que permite el rastreo de los ángulos que debe realizarse cuadro a cuadro, posicionándose en 2D sobre las articulaciones del sujeto de prueba.

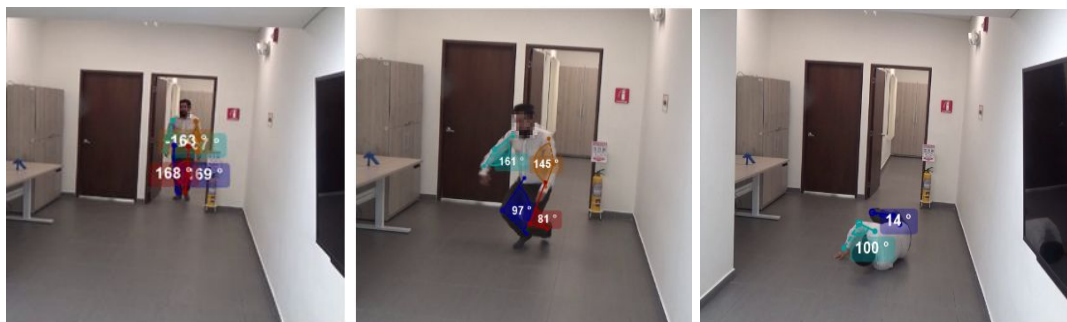


Fig. 27. En color azul claro *Rarm*, en color naranja *Larm*, en color azul oscuro *Rleg* y en color rojo *Lleg*.

5. DISCUSIÓN

Las pruebas en el laboratorio de terapia ocupacional del Rosario, presentaron algunas similitudes para los 3 tipos de formato (MTS, mp4, h.264) de video capturados con las cámaras especificadas en la Tabla II. Los videos evaluados con el sistema de detección de pose detectan la pose de personas que se encuentre inclusive hasta una distancia de aproximadamente 7 m, y presentaron fallas en algunos de los cuadros de los videos, al localizar los puntos anatómicos de la pose de una persona como se presentó en la Fig. 18b, donde las coordenadas en 'x,y' de los puntos anatómicos son iguales a 0, pero su enumeración en el mapa de puntos(Fig. 4) no es dependiente de la localización de los otros puntos.

Se observó que el modelo en algunos casos puede predecir la posición de las articulaciones en ciertos cuadros indiscriminadamente de lo que esté obstaculizando a la persona. Lo anterior se encontró al ubicar las cámaras en una misma zona a una distancia de 30 cm una de la otra y a una altura de 1,70 m aproximadamente, y se observó cómo los resultados varían de manera significativa, por lo tanto el tipo de cámara, altura y posición son factores decisivos a la hora de implementar un sistema de detección de caídas.

Posteriormente se prosiguió a procesar la información obtenida. En la prueba individual, la detección de caídas y estimación de pose fue más acertada con la cámara del teléfono inteligente Xiaomi del formato 'mp4', ya que para los videos de los formatos '.h264' y 'MTS' fallaron en la detección de la(s) persona(s) en 170 de 898 cuadros, y 124 de 978 cuadros respectivamente, antes de normalizarlos en tiempo. Lo anterior es equivalente a que el modelo falló 18,9% y 12,67% con los formatos '.h264' y 'MTS'.

A pesar de que la resolución fija es mayor en la cámara Sony y es igual a 8,24 MP, como se muestra en la tabla II de las especificaciones de las cámaras, se podría decir que este no es el factor que altera la detección de pose con el modelo utilizado, ya que la cámara del teléfono tiene una baja resolución fija y obtuvo la detección de pose en todos los cuadros del video de la prueba individual. Se hipotetiza que las características ópticas y la imagen de la cámara del teléfono inteligente marca Xiaomi deben ser un factor positivo para la detección de la pose con el modelo de detección de pose.

En la tabla VII se muestra que el error relativo más alto de los ángulos calculados fue del video capturado por la cámara del teléfono inteligente marca Xiaomi y fue de 42,87%, mientras que para el 'h264' fue de; 32,83%, y de 10,04% para el 'MTS'. Adicionalmente, se observó que la variación de los ángulos calculados por extremidad y por cada cuadro varió entre 4° y 11° para el cuadro de inicio, 58° y 79° para el cuadro del fin. Este resultado varió en gran medida para el segundo, ya que como se puede observar en los cuadros de referencia de fin del video señalados en las Fig. 18b, 19b y 20b, revelan que el modelo falló en la detección de pose del sujeto de prueba para el tipo de video de formato '.h264' capturado con el módulo de la cámara de la Raspberry, mientras que para los otros dos formatos detectó a la persona en el suelo. Estas medidas descriptivas fueron tenidas en cuenta para evaluar el sistema de caídas basado en el cálculo de los ángulos según la cámara utilizada.

Las condiciones definidas para una caída se implementaron conociendo por antelación, desde la prueba de selección de modelos, que estos fallaban cuando las personas realizaban posturas no comunes. Por lo que a manera de prevención se definió que cuando el modelo no localizaba un punto necesario para calcular un ángulo, este se equivaldría a 0.0°. Lo anterior es un factor a tener en cuenta y tiene relación con los resultados de los errores absolutos obtenidos.

Respecto a la detección de pose de la caída, realizada con los criterios expuestos, la caída que se muestra en las Fig. 18c, 19 y 20c, se detectó después de 14 y 15 cuadros, en los videos de formatos 'MTS' y 'h264' respectivamente, es decir, 0,55 y 0,62 segundos después del cuadro tomado como el real, resultando en una diferencia apenas notable, si el análisis se realizará en tiempo real.

Adicionalmente, en las gráficas de la Fig. 21 se puede observar como los ángulos varían de manera vertiginosa durante la caída que se da desde el cuadro N°50, y en mayor razón para la extremidad inferior izquierda. Intuitivamente podría decirse que no hay una correlación entre la variación de los ángulos durante las caídas además del rango articular, que es normal cuando varía entre 30° y 130° para la flexo-extensión de la rodilla y desde 20° hasta 180° para la flexoextensión del codo como indica la tabla III. Por otra parte, la desviación estándar de los tres tipos de vídeo capturado por extremidad y fue en promedio de 24.08°, lo cual corresponde a al 15,34% del rango articular del codo, y de 23,61% del rango articular de la rodilla, lo cual es un poco más que insignificante. Según estos resultados, el sistema de detección de caída debe estar basado tanto en extremidades superiores como inferiores.

Algunas variables relacionadas con el rango articular que no fueron parte del estudio de la detección de caídas, comprenden las patologías motoras que como el valgo fisiológico por ejemplo, pueden distorsionar los rangos normales de ángulos con los que se implementó el sistema de detección de caídas. Ya que esta patología comprende un rango de los ángulos negativos de 0 a -20°. La condición médica es normal en la niñez entre los 0 y 4 años, y es patológica después de esto[18].

La segunda prueba individual, permitió evaluar el desempeño del modelo cuando algunas articulaciones del sujeto de prueba se ven obstaculizadas por el uso de una silla de ruedas. En la tabla VIII se evidencia cómo esto no perjudicó la detección de pose de la persona para ninguno de los dos formatos de vídeo capturados. Un punto a tener en cuenta para la implementación del detector de caídas es que el modelo utilizado altera el resultado del cálculo de los ángulos, como se observa en la secuencia de cuadros de la Fig. 22. Si el usuario del sistema de detección de caídas presenta una discapacidad física que involucra el uso de la silla de ruedas, el sistema no se verá limitado por esta, ya que la tasa de detección de pose es un se da en todos los cuadros de manera óptima. Sin embargo, las condiciones del método de caída deberán tener en cuenta la tasa de falla de detección de predicción del modelo que varía la localización de las extremidades cuando estas se ven obstaculizadas(Tabla IX). Adicionalmente, se observó que la predicción de la localización de las articulaciones no se vio afectada si la obstaculización incluyó $\frac{1}{3}$ de la extremidad, con una parte mayor a esa el modelo no ubicó las localizaciones de los talones adecuadamente.

En la tercera prueba, a pesar de que ambos vídeos de formatos MTS y h264 detectaron la caída del sujeto de prueba, se dio cuenta de cómo el formato de video h264, capturado con el módulo de cámara de la Raspberry, resultó ser más preciso, según las condiciones de caída definida y los tiempos de inicio y fin. Además es notable que la interacción con objetos no altera la detección de caída.

La prueba con varias personas permitió evaluar la detección de pose y caída, cuando todos los sujetos del cuadro realizan diferentes acciones, e interactúan con los objetos. El modelo falla de manera considerable con la persona 3, como se muestra en la tabla XIII, en donde los puntos de las articulaciones no se localizaron de manera adecuada. Además ambos formatos de vídeo capturados, MTS y h264, detectaron una y dos caídas respectivamente de la persona 3, de manera incorrecta. La tasa de detección es baja con respecto a las otras pruebas, por lo que es necesario considerar que puede que sea necesario utilizar más de una cámara en una zona concurrida. Puede que el cálculo de los ángulos se haya visto afectado por la superposición de las

personas en más de un cuadro, por lo que la confiabilidad del sistema de detección de caídas es menor.

Para finalizar, algunos factores adicionales a tener en cuenta para implementar un sistema de detección y/o prevención de caídas con visión artificial incluyen a las pantallas que reflejan la luz, las cuales son una variable que perjudica un sistema de detección de pose que se basa en la imagen y puede alterar los resultados, haciendo que el modelo piense que existen más personas en el cuadro de las que realmente hay, generando falsas positivos y disminuyendo el nivel de confiabilidad del sistema de detección de pose. Por lo que hay que tener en cuenta adecuar las cámaras a un espacio para evitar de que esto ocurra.

Por otra parte, se realizaron los análisis de la prueba individual utilizando el *software* Kinovea de análisis deportivo, los cuales se realizan ubicando de manera manual los puntos de referencia de las articulaciones, y se debe corregir su ubicación en cada uno de los cuadros que componen el video. Este método es poco preciso, tedioso y demanda bastante tiempo para verificar la ubicación de los puntos a medida que transcurre el video. Además requiere de un paso de calibración del sistema que incluye la configuración de medidas antropométricas, como la altura y otras medidas corporales, que claramente variará según el sujeto presente en los cuadros del video. Otra desventaja de este método, es que el análisis se limita a una sola persona en el video. El análisis del video de la prueba individual de la caída simulada supuso 3.5h, mientras que con la estimación de pose y el sistema de detección de caída el análisis del mismo video se realiza en 20 minutos o menos.

6. CONCLUSIONES

El desarrollo del sistema de detección de pose y de caídas por medio de VA con las cámaras de diferentes rangos y resoluciones, arrojó que es posible implementar un sistema confiable para la detección de caídas con cámaras de bajo costo en un ambiente de la vida cotidiana. Se comprobó con el estudio que el desempeño del modelo sí varía según el tipo de dispositivo que se utilice y varió notablemente con el módulo de la cámara de la Raspberry, por lo que el tipo de cámara a utilizar a la hora de implementar un sistema de detección de caídas es un factor de gran relevancia.

Intuitivamente se podría afirmar que la resolución de imagen del teléfono inteligente es el gran factor de la detección de pose exitosa que le permitió obtener un nivel de confianza a la par de los otros dos tipos de cámaras para la detección en los cuadros de los videos evaluados.

Los videos capturados con el teléfono inteligente y con el módulo de la cámara de la Raspberry, lograron captar la caída con una variación aproximada al tercer formato capturado con la cámara Sony, por lo que es factible implementar con cámaras de gama media un sistema con una buena relación costo-beneficio de la detección de pose y caídas.

Además, se dio cuenta que la obstaculización de las partes del cuerpo no es un impedimento para este tipo de sistemas de detección de caídas, por lo que no existe una limitación para las personas que hacen uso de artefactos extracorpóreos, como la silla de ruedas, o bastones, por lo que es un sistema fácilmente adaptable a un ambiente clínico o residencial.

Para finalizar, de las grandes ventajas del método propuesto con VA, es que este método permite simplificar la captura de videos para la detección de pose y caídas, debido a que el equipo adicional, de sensores y cámaras de alto costo es innecesario, disminuyendo no solo el tiempo de la toma de capturas si no el procesamiento de la información que se genera. Lo anterior, evitará el tedioso proceso de utilizar equipo costoso que puede dificultar el objeto de su propio uso, y permitirá realizar esta función de manera rápida y eficaz, despreciando variables antropométricas.

7. RECOMENDACIONES Y TRABAJOS FUTUROS

Teniendo en cuenta la comparación de la detección de pose realizada en este estudio, y utilizado diferentes dispositivos de resolución similar, se obtuvo que es posible un resultado aproximado con un teléfono inteligente de gama media y con una tarjeta de desarrollo, para generarse un sistema de detección de caídas básico y preciso. Debido al limitado tiempo presencial para el uso de *hardware* y de un laboratorio de marcha, y a los imprevistos generados por la pandemia, se recomienda validar el desarrollo en un laboratorio de marcha; e implementar los métodos desarrollados en este estudio en un dispositivo como la TPU Google Coral para análisis remoto en tiempo real.

A continuación se listan algunos trabajos futuros que se podrían continuar desarrollando:

- Analizar el impacto de las discapacidades motoras que pueden introducir nuevas variables y rangos de movimiento mayores o menores, que no se incluyeron en el sistema de detección de caídas implementado en este estudio.
- Agregar un sistema de aprendizaje automático que comprenda la acción de las personas basado en el cálculo de los ángulos o calculando el centro de masa de las personas presentes en el video.
- Complementar con un sistema estéreo de varias cámaras de bajo costo, localizadas en zonas opuestas de un espacio, que permitan realizar una detección de pose incluso en 3D para aumentar la exactitud y precisión de la detección de caídas.
- Analizar la posibilidad de un sistema para la detección de caídas a bajo costo, debido a que con una raspberry no se requiere elementos costosos para su implementación en tiempo real. El sistema podría ser accedido de forma remota, sería adaptable a cualquier ambiente, y permitiría un control a distancia y portable, para prever caídas que podrían agravar la salud, con fracturas de cadera y/o rodillas, hematomas generales, y demás, que puedan vulnerar y debilitar el estado de personas vulnerables.

REFERENCIAS

- [1] Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
- [2] López Quintero, M. I. (2016). Estimación de la pose humana 2D en imágenes estéreo.
- [3] Li, Xiu, et al. 3D Pose Detection of Closely Interactive Humans Using Multi-View Cameras. *Sensors*, 2019, vol. 19, no 12, p. 2831.
- [4] Wang, C., Wang, Y., & Yuille, A. L. (2013). An approach to pose-based action recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 915-922).
- [5] Andriluka, M., Pishchulin, L., Gehler, P., & Schiele, B. (2014). 2d human pose estimation: New benchmark and state of the art analysis. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3686-3693).
- [6] Google. Introduction to TensorFlow. Disponible: [tensorflow.org/learn](https://www.tensorflow.org/learn)
- [7] H.Hayes.(2017,Marzo) Pelvis. Disponible: [http://c-motion.com/v3dwiki/index.php?title=Helen_Hayes_\(Davis\)_Pelvis](http://c-motion.com/v3dwiki/index.php?title=Helen_Hayes_(Davis)_Pelvis)
- [8] O'Shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*.
- [9] Leitch, J., Stebbins, J., Paolini, G., & Zavatsky, A. B. (2011). Identifying gait events without a force plate during running: A comparison of methods. *Gait & Posture*, 33(1), 130-132.
- [10] Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., ... & Klingner, J. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. *arXiv preprint arXiv:1609.08144*.
- [11] Tang, L., Garvin, M. K., Lee, K., Alward, W. L., Kwon, Y. H., & Abramoff, M. D. (2011). Robust multiscale stereo matching from fundus images with radiometric differences. *IEEE transactions on pattern analysis and machine intelligence*, 33(11), 2245-2258.
- [12] Gómez-González, E., Gomez, E., Márquez-Rivas, J., Guerrero-Claro, M., Fernández-Lizaranzu, I., Relimpio-López, M. I., ... & Capitán-Morales, L. (2020). Artificial intelligence in medicine and healthcare: a review and classification of current and near-future applications and their ethical and social impact. *arXiv preprint arXiv:2001.09778*.
- [13] LeCun, Y. (2018). The Power and Limits of Deep Learning: In his IRI Medal address, Yann LeCun maps the development of machine learning techniques and suggests what the future may hold. *Research-Technology Management*, 61(6), 22-27.
- [14] ART: Advanced Realtime Tracking. Motion Capture. Abril 10, 2020.[Online]. Disponible: <https://ar-tracking.com/applications/motion-capture/>
- [15] Cao, Z., Hidalgo, G., Simon, T., Wei, S. E., & Sheikh, Y. (2018). OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields. *arXiv preprint arXiv:1812.08008*.
- [16] LeCun, Y., & Bengio, Y. (1995). Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, 3361(10), 1995.
- [17] Kim, I. Julio, 2019. Deep Pose Estimation implemented using Tensorflow with Custom Architectures for fast inference.[Online]. Disponible: <https://github.com/ildoonet/tf-pose-estimation>
- [18] Valores normales para la amplitud de movimiento de las articulaciones. Manual MSD: versión para profesionales.[Online]. Disponible: https://www.msdmanuals.com/es/professional/multimedia/table/v1128315_es
- [19] Gómez, J. A. V. (2016). Problemas bioéticos emergentes de la inteligencia artificial. *Diversitas: perspectivas en psicología*, 12(1), 137-147.
- [20] Piñeros Lourenco, H. F. *Evaluación del desarrollo y viabilidad de un sistema de detección de movimiento aplicable a pacientes de alto riesgo en caídas* (Universidad del Rosario).

ANEXOS

ANEXO-A

| Actividad N° | Inicio | Fin | 21/01/2020 | 25/01/2020 | 28/01/2020 | 1/02/2020 | 4/02/2020 | 8/02/2020 | 11/02/2020 | 15/02/2020 | 18/02/2020 | 24/02/2020 | 25/02/2020 | 29/02/2020 | 3/03/2020 | 10/03/2020 | 11/03/2020 | 12/03/2020 | 14/03/2020 | 17/03/2020 | 21/03/2020 | 24/03/2020 | 28/03/2020 | 29/03/2020 | 30/03/2020 | 31/03/2020 | 4/04/2020 | 7/04/2020 | 11/04/2020 | 14/04/2020 | 18/04/2020 | 21/04/2020 | 25/04/2020 | 28/04/2020 | 2/05/2020 | 3/05/2020 | 22/05/2020 | | | | | | | |
|--------------|------------|------------|------------|------------|------------|-----------|-----------|-----------|------------|------------|------------|------------|------------|------------|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|-----------|-----------|------------|------------|------------|------------|------------|------------|-----------|-----------|------------|--|--|--|--|--|--|--|
| | | | 1 | 21/01/2020 | 8/02/2020 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2 | 5/02/2020 | 19/02/2020 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3 | 24/02/2020 | 24/02/2020 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 4 | 24/02/2020 | 9/03/2020 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 5 | 11/03/2020 | 12/03/2020 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 6 | 30/03/2020 | 30/03/2020 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 7 | 24/03/2020 | 11/04/2020 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 8 | 14/04/2020 | 24/04/2020 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 9 | 21/01/2020 | 2/05/2020 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 10 | 24/04/2020 | 24/04/2020 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 11 | 3/05/2020 | 3/05/2020 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 12 | 22/05/2020 | 22/05/2020 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

ANEXO-B

ESTUDIO DE INVESTIGACIÓN “Herramienta automática para análisis de actividades cotidianas en un ambiente residencial a través de técnicas de inteligencia artificial”

DOCUMENTO DE CONSENTIMIENTO INFORMADO

- Por favor, lea cuidadosamente esta información sobre el estudio de investigación titulado **“Herramienta automática para análisis de actividades cotidianas en un ambiente residencial a través de técnicas de inteligencia artificial”**.
- Siéntase en completa libertad de preguntar al personal del estudio todo aquello que no entienda.
- Una vez haya comprendido la información, se le preguntará si desea participar del estudio. En caso afirmativo, deberá firmar este documento y recibirá una copia.

DESCRIPCIÓN GENERAL

Actualmente, las estadísticas en nuestro país reflejan una tendencia incremental de la esperanza de vida de las personas mayores. Esto conlleva a nuevos retos que incluyen el mantenimiento de la funcionalidad y una vida independiente, siendo factores fundamentales para una mejor calidad de vida. Para la persona mayor es de gran importancia mantener la independencia y poder envejecer en el hogar, teniendo en cuenta que es el entorno central donde transcurre su vida. Estas características de calidad de vida son aplicables también a personas en condición de discapacidad, las cuáles ven como en muchos casos, su independencia se ve amenazada, sobretodo en tareas cotidianas en sus hogares.

Los sistemas automáticos apoyados por análisis de datos que puedan provenir de diferentes fuentes como sensores y video soportan nuevas alternativas que pueden emplear personas dedicadas al cuidado de las poblaciones mencionadas anteriormente. De esta forma, los individuos bajo cuidado pueden vivir de una forma más independiente y permite reducir la carga física, emocional y ocupacional de quienes los cuidan.

El objetivo de este proyecto es proponer un sistema automático que permita discriminar las actividades de la vida diaria que realiza un sujeto en un ambiente controlado residencial, a partir de datos de sensores y registros de video, empleando inteligencia artificial.

OBJETIVO DEL ESTUDIO

Desarrollar un sistema automático de reconocimiento de actividades cotidianas a partir del uso de información de sensores y video, y técnicas de la inteligencia artificial en un ambiente controlado.

¿POR QUÉ FUE USTED ELEGIDO PARA PARTICIPAR EN ESTE ESTUDIO?

Usted fue elegido para este estudio por ser parte de la planta profesoral de los programas de ingeniería biomédica y terapia ocupacional de la Universidad del Rosario.

RIESGOS Y BENEFICIOS

Este estudio tiene un riesgo mínimo porque se trabajarán únicamente datos provenientes de sensores que son colocados en forma de manilla en los sujetos participantes, y de registros de video, donde se anonimizará el sujeto en cuestión. Dentro de los beneficios estará el aporte a la generación de una herramienta automática para identificación de actividades cotidianas que se puede emplear posteriormente en otros estudios que ayuden a poblaciones vulnerables. Así mismo, la Universidad del Rosario en sus programas de ingeniería biomédica y terapia ocupacional con los resultados obtenidos en este proyecto, como visibilidad nacional e internacional de los grupos de investigación involucrados.

¿COMO SERÁ LA PARTICIPACIÓN EN EL ESTUDIO?

Su participación requiere de los siguientes procedimientos, que usted podrá libremente aceptar o rechazar:

1. El lugar para realizar estos procedimientos será en la casa de actividades de la vida diaria, que la Universidad del Rosario tiene en la sede de Quinta Mutis. Es un ambiente seguro y tranquilo, donde previamente han sido instaladas unas cámaras e infraestructura necesaria para captura de datos.
2. Se le realizarán unas preguntas relacionadas con su edad, sexo, peso y estatura para mantener registro por parte de alguno de los investigadores
3. Se le colocará una manilla, no más grande que un reloj común, y se le pedirá que en 30 minutos simule actividades comunes que realice en su hogar como sentarse a ver televisión, leer un libro, ir a la cocina, simular cocinar, dormir, hablar por teléfono, por ejemplo, etc. Mientras realiza dichas actividades, se guardará registro de los datos que producen los sensores en su manilla y registro de video de dichas actividades.
4. Posteriormente, se almacenarán dichos datos con un identificador interno, de tal manera que los investigadores puedan guardar fielmente dicha información.

GARANTÍAS DE SU PARTICIPACIÓN

Su participación será completamente voluntaria y tendrá el derecho de retirarse en cualquier momento del estudio si usted así lo desea. Igualmente, si en algún momento desea que la información que usted brinda no sea utilizada por los investigadores, lo podrá comunicar y respetaremos su decisión. Para su seguridad, toda la información que sea recolectada en este estudio será confidencial y anónima. Para propósitos de análisis de datos se le asignará un código, por lo tanto, sus datos personales no serán utilizados en ningún momento. También, la información

personal no será vinculada a los datos. Los datos obtenidos serán utilizados para entrenamiento de algoritmos de inteligencia artificial, que eventualmente y de manera anónima, se usarán en presentaciones de investigación, publicaciones académicas y reportes, únicamente si usted está de acuerdo con ello.

La información se mantendrá por un mínimo de 5 años después de que el estudio se haya completado, siguiendo la política de la Universidad. Después de este tiempo la información será destruida y únicamente los resultados finales se mantendrán almacenados. Todos los archivos se guardarán en forma digital en un computador dedicado al proyecto, protegido por contraseñas y encriptación de datos.

Finalmente, se informa que participar en el estudio no tiene ningún costo. Los procedimientos y exámenes que se le practiquen en este estudio tampoco tendrán costo. Ni usted, ni otra persona involucrada en el estudio, recibirá beneficios políticos, económicos o laborales como compensación por su participación.

MANEJO DE LOS DATOS DE INVESTIGACIÓN

La información se mantendrá bajo estricta confidencialidad y no se utilizará su nombre o cualquier otra información que pueda identificarlo personalmente.

Toda la información que se obtenga de este estudio de investigación se utilizará únicamente con el propósito que aquí se comenta. Los investigadores de este estudio son los únicos autorizados para acceder a los datos que usted suministre. Para esto, los datos obtenidos solo estarán almacenados en un computador dedicado para tal fin en la Universidad del Rosario, guardados bajo contraseña y encriptación por seguridad. Posterior a cinco (5) años de la fecha de la adquisición de los datos, ésta se eliminará según políticas de la Universidad del Rosario.

Le informaremos de los resultados obtenidos en el estudio. También podrá contactar al personal del estudio e informarnos cualquier situación anormal o inesperada en cualquier momento.

Cualquier duda sobre los datos almacenados, éstos solo serán manipulados y almacenados por los investigadores del proyecto. Consulte al final de este documento para saber el responsable como investigador principal.

ACEPTACIÓN

Por favor marque con una “X” en caso que acepte o no acepte lo siguiente:

| | | |
|--|---------------|------------------|
| Autorizo a los investigadores del estudio “Herramienta automática para análisis de actividades cotidianas en un ambiente residencial a través de técnicas de inteligencia artificial” para: | Acepto | No acepto |
|--|---------------|------------------|

| | | |
|--|--|--|
| · Realizar los procedimientos descritos en este documento, necesarios para la realización del estudio de investigación | | |
| · Proporcionar información sobre mi edad, sexo, peso y altura | | |
| · Colocar una manilla en cualquiera de mis brazos para registrar información de sensores inerciales | | |
| · Hacer grabaciones en video | | |
| · Almacenar datos sobre actividades cotidianas que simule en el lugar del experimento | | |
| · Comunicarse conmigo para hacer los seguimientos requeridos por el estudio | | |
| · Comunicarse conmigo para invitarme a participar de otros estudios de investigación | | |

Participante

Nombre

Cédula

Firma

Día/Mes/Año

ESPACIO RESERVADO PARA EL INVESTIGADOR

En nombre del estudio “**Herramienta automática para análisis de actividades cotidianas en un ambiente residencial a través de técnicas de inteligencia artificial**”, me comprometo a guardar la identidad de _____ como participante. Acepto su derecho a conocer el resultado de todas las pruebas realizadas y a retirarse del estudio a su voluntad en cualquier momento. Me comprometo a manejar los resultados de esta evaluación de acuerdo a las normas para la realización de investigación en Colombia (Resolución 8430 de 1993 y Resolución 2378 de 2008) y la ley para la protección de datos personales (Ley estatutaria 1581 de 2012).

Nombre: _____

Documento de Identidad No. _____

Firma: _____

Fecha (día/mes/año) ____/____/____

¿INFORMACIÓN O PREGUNTAS ADICIONALES?

Si en algún momento desea obtener información adicional sobre el estudio puede contactar a:

Alvaro David Orjuela Cañón,

Investigador principal.

Escuela de Medicina y Ciencias de la Salud.

Universidad del Rosario

Carrera 24 # 63C – 69. Edificio administrativo de la EMCS en Bogotá D.C. – Colombia,

Teléfono: 2970200 ext. 3479.

Correo electrónico: alvaro.orjuela@urosario.edu.co

ANEXO-C

Comité de Ética de la Universidad del Rosario

COMITÉ DE ÉTICA EN INVESTIGACIÓN DE LA UNIVERSIDAD DEL ROSARIO
SALA DE CIENCIAS DE LA VIDA

MIEMBROS

CARLOS ENRIQUE TRILLOS, PRESIDENTE
MÉDICO CIRUJANO; MSc. EPIDEMIOLOGÍA

JUAN GUILLERMO PÉREZ CARREÑO
MÉDICO BIOCITISTA;
SECRETARIO TÉCNICO

RAMON FAYAD NAFFAH
Lic. EDUCACIÓN; PHD MATEMÁTICAS

CARLOS ALBERTO CALDERÓN
MÉDICO; MSc. FARMACOLOGÍA

CARLOS GUILLERMO CASTRO
ABOGADO; ESP. DERECHO MÉDICO SANITARIO

GLEIDY VANESSA ESPITIA
FONOAUDIÓLOGA, MSc Y PHD EN CIENCIAS

JAD CONSTANZA ROJAS
BACTERIÓLOGA; MSc. SALUD PÚBLICA;
MSc. BIOÉTICA

KATHERIN QUINTERO PARRA
QUÍMICA FARMACÉUTICA

LUISA FERNANDA RAMÍREZ
PSICÓLOGA- PHD PSICOLOGÍA SOCIAL Y DE LA SALUD

MAGDA MILENA GAVIRIA
BIÓLOGA; MSc EN BIOLOGÍA

MARTHA ROCÍO TORRES NARVÁEZ
FISIOTERAPEUTA, MSc BIOÉTICA

ANDREA TORRES RUIZ
INGENIERA BIOMÉDICA; MSc ING. ELECTRÓNICA Y DE
COMPUTADORES

CESAR PAYAN GÓMEZ
MÉDICO, MSc GENÉTICA HUMANA; PHD CIENCIAS
BÁSICAS

JESUS ORLANDO CROCE
ABOGADO

KAREN JULIETH MORENO
FISIOTERAPEUTA; ESP. EPIDEMIOLOGIA

MARTHA ISABEL BAUTISTA DUEÑAS
AUXILIAR ADMINISTRATIVA

DVO005 926-CV1131

Bogotá D. C., 09 de octubre de 2019

Doctor

ALVARO DAVID ORJUELA CAÑÓN

Investigador Principal

Estudio: **"Herramienta automática para análisis de actividades cotidianas en un ambiente residencial a través de técnicas de inteligencia artificial"**.

Bogotá, D. C.

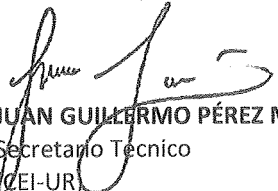
Respetado Investigador:

El Comité de Ética en investigación de la Universidad del Rosario (CEI-UR) evaluó de forma expedita la segunda versión su proyecto de investigación "Herramienta automática para análisis de actividades cotidianas en un ambiente residencial a través de técnicas de inteligencia artificial". Fecha de recepción: 16 de septiembre de 2019.

Luego de haber tenido en cuenta las observaciones efectuadas por el CEI-UR, se APRUEBA el protocolo en referencia junto con la información anexa.

Para el comité de ética es importante acompañarlo durante la ejecución del estudio. Por favor no dude en contactarnos en caso de tener alguna inquietud o de necesitar apoyo para el análisis de alguna situación específica. De igual forma le recomendamos notificar cualquier modificación en la ejecución del estudio no expuesta en la aprobación inicial del proyecto.

Cordialmente,


JUAN GUILLERMO PÉREZ MD, MSc
Secretario Técnico
(CEI-UR)



c.c. Archivo
Proyectó: Martha Isabel Bautista

Este comité se rige por los lineamientos jurídicos y éticos del país a través de las resoluciones 8430 de 1993 y 2378 de 2008 del Ministerio de Salud y Protección Social. Igualmente, se siguen los acuerdos contemplados en la declaración de Helsinki (Fortaleza, Brasil 2013) y de la Conferencia Internacional de Armonización para las Buenas Prácticas Clínicas. Recuerde visitar nuestra página web, en donde encontrará información actualizada de los procedimientos del Comité de Ética en Investigación de la Universidad del Rosario, así como cursos en ética de la investigación de acceso libre.
<https://www.urosario.edu.co/investigacion/Soporte-a-la-Investigacion/Comite-de-etica-en-investigacion/#id-tabs>



DVO005 926-CV1131

Página 1 de 1