

# Dispositivo de Evaluación de Poses Integrado a una Arquitectura Terapéutica IoT

Francisco José Moreno Alvarez, Walter Hernández  
[francisco.j.moreno@ucv.ve](mailto:francisco.j.moreno@ucv.ve), [walter.hernandez@ciens.ucv.ve](mailto:walter.hernandez@ciens.ucv.ve)

Centro de Computación Gráfica, Universidad Central de Venezuela, Facultad de Ciencias, Caracas, Venezuela

**Resumen:** En el presente trabajo se describe el proceso de desarrollo e integración de un dispositivo inteligente a una arquitectura basada en *IoT* orientada a la terapia física. El dispositivo inteligente corresponde con un sistema que monitorea la postura del paciente mientras este realiza un conjunto de posturas propuestas por un especialista médico. El dispositivo procura ser de bajo costo monetario y no intrusivo, permitiéndole al paciente tener una libre movilidad durante el ejercicio. Para su desarrollo se proponen un conjunto de alternativas acerca del módulo encargado de la adquisición de las características del paciente (esqueleto). Como resultado se obtiene un sistema virtualmente agnóstico al *hardware* utilizado para la captura del esqueleto del paciente. De igual manera se proponen distintos mecanismos para comparar la postura del paciente contra la esperada, procurando realizar una detección eficaz y proveer de herramientas que le permitan tanto al paciente como al personal médico evaluar su rendimiento. Por otra parte, se indican las consideraciones tomadas al momento del desarrollo del dispositivo para poder integrarlo a la arquitectura *IoT* previamente implementada. Se describe el proceso para registrar los parámetros requeridos para configurar el dispositivo desde una herramienta *web*, el proceso para asignar un ejercicio a un paciente a través del dispositivo desarrollado, así como las consideraciones tomadas para evaluar los resultados provistos por el mismo. Finalmente, se comprueba la eficacia del método de comparación de poses, demostrando ser altamente efectivo independientemente de la posición y estatura del usuario. De igual manera, se compara la eficiencia de los mecanismos de adquisición planteados, proporcionando una alternativa capaz de operar inclusive en dispositivos móviles.

**Palabras Clave:** Poses; *IoT*; Terapia; Dispositivo Inteligente; Internet de las Cosas; Rehabilitación.

**Abstract:** In this work the development and integration of a smart device to a physical therapy *IoT* based architecture is presented. The smart device, supervises the patient while performing a set of postures configured by a medical specialist. The proposed device is designed to have a low monetary cost and to be unobtrusive throughout the therapy's execution, allowing the patient to move freely and with no physical restrictions. Several alternatives are studied to handle the extraction of the user posture defining characteristics, such as the skeleton. As a result, the proposed device is considered virtually agnostic to the *hardware* used to capture the user's skeleton. Furthermore, several techniques to compare the user's posture with the target pose are studied, pursuing an effective method which provides the tools required from the patient and therapist to qualify the therapy's execution performance. On the other hand, a set of considerations are taken into account to allow an easy integration of the device with an existing implementation of an *IoT* architecture. The process to register all the required device parameters through a web platform, including the steps to assign an exercise based on the smart device and all the requirements to evaluate the results of each exercise are also described. Finally, the effectiveness of the proposed technique to compare the poses is tested, proving to be highly effective independently of the location and height of the patient. Likewise, the efficiency of the skeleton acquisition techniques are demonstrated, providing an alternative capable of working even on mobile devices.

**Keywords:** Poses; *IoT*; Therapy; Smart Device; Internet of Things; Rehabilitation.

## I. INTRODUCCIÓN

Se estima que globalmente alrededor de 2.4 billones de personas conviven con algún tipo de afectación de salud que podría mejorarse directamente a través de la rehabilitación. Sin embargo, actualmente estas necesidades están siendo cubiertas en un porcentaje muy bajo. En países de mediano a bajo

ingreso por ejemplo, más de la mitad de la población que requiere este tipo de servicio no lo recibe. Por otra parte los servicios de rehabilitación han detenido abruptamente su funcionamiento en aproximadamente un 60-70% de los casos, debido a las condiciones generadas por eventos como la pandemia por COVID-19 desde principios del año 2020 [1].

Así mismo, múltiples investigaciones [2][3] confirman la efectividad del ejercicio físico como una herramienta para la recuperación y prevención de ataques cerebro vasculares, e indican su capacidad para actuar como un agente de retardo de las enfermedades asociadas al envejecimiento. Por ende, es necesario que la población mantenga un estilo de vida activo y realice ejercicio. Dichos ejercicios deben ser diseñados y evaluados por personal especializado en las condiciones particulares del paciente. Dados los problemas de logística, infraestructura y capacidad monetaria que involucran prestar y recibir sesiones de ejercicio o terapia de manera presencial, se propuso una arquitectura basada en *IoT* [4] que permita al paciente realizar sesiones de terapia de manera remota y, al personal médico evaluar su progreso. Dicha arquitectura pretende aliviar la precariedad del sistema de salud, que se agrava aún más por situaciones extraordinarias como la mencionada pandemia del COVID-19.

En el trabajo actual se presenta un dispositivo terapéutico que forma parte de la arquitectura basada en *IoT* y que tiene como objetivo fortalecer la flexibilidad y movilidad de los pacientes que lo utilicen. Dicho dispositivo solo requiere el uso de un computador tradicional y una cámara *web* (en esta primera fase), haciéndolo de fácil acceso y sin requerir un gasto económico adicional para el paciente o el centro hospitalario que lo provea.

El resto del trabajo se presenta en 4 secciones. La Sección II describe brevemente algunos trabajos relacionados a la propuesta que se presenta en la Sección III. La Sección IV plantea las pruebas que se realizaron para evaluar su capacidad y efectividad de evaluar la calidad del ejercicio realizado por el paciente. Finalmente, la Sección V expone las conclusiones del trabajo y trabajos futuros.

## II. TRABAJOS RELACIONADOS

Ahmad et al. [5] presentan un robot de asistencia social que motiva a personas de la tercera edad a realizar ejercicios de movilidad recomendados por el servicio nacional de salud de Inglaterra [6]. Mientras la persona realiza los ejercicios que el robot le propone, este evalúa que se realicen de la manera correcta y proporciona instrucciones y palabras de aliento al paciente para que complete el ejercicio.

Los movimientos del paciente son capturados mediante el uso de un *Microsoft Kinect* que genera la representación de su esqueleto. La postura registrada se evalúa mediante un modelo de cinemática angular que compara los ángulos de los huesos del paciente con un conjunto de parámetros preestablecidos. El modelo se complementa con zonas espaciales donde deben encontrarse las articulaciones del esqueleto registrado del paciente.

El sistema es probado con un conjunto de pacientes de la tercera edad, que manifestaron su satisfacción de contar con un asistente personal (el robot de asistencia social) que les proponga ejercicios y los motive durante su ejecución. En la Figura 1, se puede observar a un paciente siendo asistido por el robot propuesto en la investigación.



Figura 1: Paciente Utilizando el Sistema Propuesto por [5]

Por otra parte, Ian et al. [7] indican la importancia de fortalecer y mantener el tono muscular del cuerpo para la salud humana. Hacen énfasis en el Yoga como una herramienta para fortalecer la musculatura y mejorar la flexibilidad del prácticamente. Sin embargo, mencionan que el proceso de aprendizaje del Yoga tiende a ser entorpecido por los errores de juicio de los distintos instructores de Yoga. Por ende, proponen un sistema para aprender a realizar posturas de Yoga mediante el uso del computador.

En este sistema, el aprendiz es grabado por una cámara mientras realiza una postura siguiendo un conjunto de instrucciones. Las imágenes capturadas, se envían a un computador que extrae el esqueleto del usuario mediante una red neuronal denominada *OpenPose*<sup>1</sup>. Posteriormente, evalúa los ángulos de los huesos del estudiante mediante un modelo angular similar a [5]. En caso de que el estudiante realice correctamente la postura, su esqueleto es desplegado en color verde y en caso contrario en color rojo. El sistema se complementa utilizando una tabla (*hardware especializado*) que evalúa el balance del paciente. En la Figura 2 se puede observar a un practicante haciendo uso del sistema.

Finalmente Santosh et al. [8], proponen un sistema para aprender y reconocer posturas de Yoga (*asanas*) exclusivamente mediante el uso de técnicas basadas en aprendizaje profundo. En este caso, utilizan *OpenPose* al igual que [7] para identificar los puntos que correspondan con las articulaciones del usuario en una imagen. La red previa alimenta a una red neural recurrente con memoria a corto y largo plazo, que relaciona las posturas con información temporal para aprender la secuencia de movimientos que conlleva realizar correctamente una *asana*.

Los datos de entrenamiento para las redes, fueron creados a partir de vídeos de 15 individuos realizando *asanas*, y se probó con 12 individuos distintos a los de la fase de entrenamiento. Como resultado, se obtuvo un 90% de exactitud en la pruebas realizadas. En la Figura 3 se puede observar a un usuario haciendo uso del sistema propuesto.

En el sistema desarrollado en esta investigación se propone una alternativa al método de comparación angular utilizado

<sup>1</sup><https://github.com/CMU-Perceptual-Computing-Lab/openpose>



Figura 2: Usuario Utilizando el Sistema Propuesto por [7]

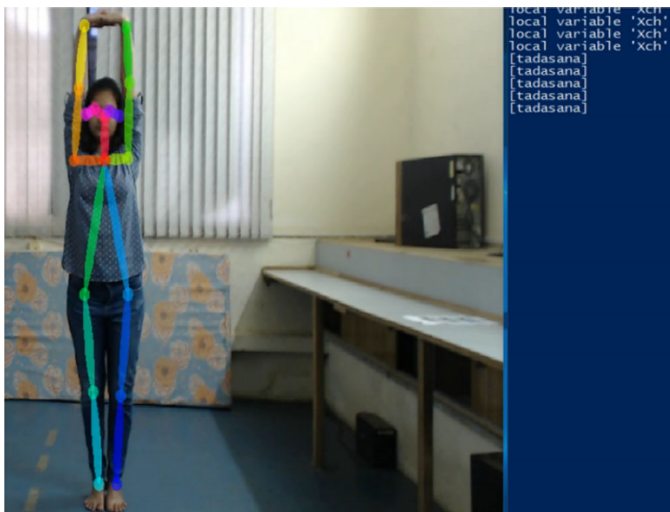


Figura 3: Usuario Utilizando el Sistema Propuesto por [8]

en [5] y [7], capaz de discernir entre poses distintas cuyos ángulos internos son los mismos, sin la necesidad de requerir información adicional. Se experimenta con distintos dispositivos de adquisición del esqueleto y se compara su eficiencia. Finalmente, se acopla la propuesta a la red de dispositivos terapéuticos que forman parte de la arquitectura basada en *IoT* producto de una investigación previa [4].

### III. PROPUESTA

El dispositivo que se propone a continuación, tiene como objetivo evaluar el rango de movilidad del paciente mientras imita una postura propuesta por el personal médico. Para alcanzar la meta es necesario definir:

- Un mecanismo para capturar la postura actual del paciente.

- Un algoritmo que permita evaluar la similitud entre la postura del paciente y la propuesta por el personal médico.
- Herramientas que informen al paciente el grado de similitud contra la pose objetivo, para facilitar la tarea de imitarla.
- Elementos que le permitan al médico visualizar y diseñar la postura para el paciente y a este último, tener un punto de referencia de la postura a seguir.
- Finalmente, una aplicación que le permita al médico configurar la pose y su homólogo que evalúe la postura del paciente.

Es importante destacar la necesidad de desarrollar un dispositivo de fácil acceso, por ende, debe ser de bajo costo monetario y en la medida de lo posible no depender de *hardware* especializado que involucre un gasto adicional para el paciente o el centro hospitalario que lo provea.

#### A. Adquisición de Postura

Para identificar la postura del paciente es necesario conocer la posición en el espacio de cada una de las articulaciones que componen su esqueleto. Además, ya que la aplicación será utilizada por personas con movilidad reducida, es importante que el mecanismo de adquisición no dificulte la ya mermada movilidad del paciente. Dada estas observaciones, se experimentó con tres alternativas de adquisición no invasivas, que infieren la posición de las articulaciones del paciente a partir del análisis de imágenes.

La primera alternativa fue el *Microsoft Kinect*. Dispositivo especialmente desarrollado para la adquisición del esqueleto de múltiples usuarios dentro de su rango de visión [9]. Fue concebido originalmente para ser utilizado en conjunción con la consola de videojuegos *Xbox 360*, con la intención de desarrollar aplicaciones interactivas. Utiliza imágenes de profundidad (por láser), haciéndolo poco susceptible a las características de iluminación del entorno en el que se encuentre el usuario y, permitiéndole inferir la posición en tres dimensiones de 20 articulaciones sobre el mismo. La clara desventaja de trabajar con esta herramienta, es que el paciente o centro hospitalario debe contar con el *hardware* especializado *Microsoft Kinect*.

*OpenPose* surge como alternativa al no requerir de un dispositivo especializado de adquisición. Utiliza como insumo imágenes que pueden ser adquiridas a través de cualquier cámara convencional. *OpenPose* es una librería de *software* libre basada en aprendizaje profundo [10][11][12][13], capaz de extraer la posición 143 puntos claves por cada usuario presente en una imagen. Dichos puntos se descomponen en: 25 puntos generales para el cuerpo del usuario, 6 puntos que identifican cada parte del pie, 21 puntos por mano y 71 para la cara. Esta alternativa presenta un claro avance en precisión y cantidad de información registrada contra la previa. Sin embargo, requiere de un computador con alto poder de cómputo para alcanzar velocidades de procesamiento que se consideren interactivas. La librería cuenta con un apartado que permite la aceleración del proceso de clasificación mediante la tarjeta gráfica, pero es dependiente de la arquitectura propietaria *CUDA* del fabricante

de *hardware Nvidia*. Si bien es cierto que el trabajar con imágenes tradicionales presenta una ventaja ante la facilidad de acceso a la tecnología, trae como consecuencia que los puntos claves extraídos por la librería se encuentren ubicados en el espacio en dos dimensiones de la imagen.

Finalmente, *PoseNet* al igual que *OpenPose* es una librería basada en aprendizaje profundo [14][15] propuesta por *Google* y es capaz de extraer la posición 17 puntos claves en el cuerpo de cada usuario presente en una imagen. Para realizar la extracción de los mismos puede hacer uso de un modelo de clasificación diseñado para trabajar en dispositivos móviles y de bajo poder de cómputo [16] o un modelo de clasificación más preciso [17] pero más lento. Adicionalmente, *PoseNet* tiene la capacidad de ejecutarse directamente en un navegador *web*, al ser soportado por la versión *web* de *TensorFlow*<sup>2</sup>. Estas características hacen que sea un fuerte competidor contra las alternativas previamente mencionadas, ya que a pesar de ser menos preciso que *OpenPose*, tiene la capacidad de operar en computadoras convencionales a tasas interactivas y no requerir *hardware* especial para la adquisición de las imágenes a diferencia del *Microsoft Kinect*, por lo que se considera como la alternativa más accesible de las tres.

A pesar de que *PoseNet* se presenta como la mejor alternativa en lo que confiere al tema de accesibilidad, el sistema fue desarrollado para que pudiera operar independientemente del método de adquisición seleccionado. Por ende, se desarrolló un esqueleto estándar que será utilizado en las siguientes etapas del sistema de evaluación. El esqueleto estándar esta compuesto por 14 puntos claves que pueden ser extraídos de los esqueletos que proveen las tres alternativas y, en el caso específico de *PoseNet*, el punto clave correspondiente al inicio del cuello puede ser inferido como el punto medio entre los hombros del paciente. El esqueleto estándar puede ser visualizado en la Figura 4. Los colores utilizados para representar cada hueso fueron elegidos de una paleta de 21 colores de alto contraste, diseñada para facilitar su identificación y memorización por parte del usuario [18].

### B. Algoritmo de Evaluación

Posterior a la adquisición de los puntos claves que componen el esqueleto del paciente, la postura que describa debe ser comparada contra la pose propuesta por el personal médico. Dicha comparación debe:

- Ser independiente de la posición del paciente en el espacio.
- Ser invariante ante la posible diferencia en estaturas de los pacientes.
- Permitir algún mecanismo de apoyo visual que guíe al paciente para alcanzar la pose deseada.

Una forma de comparar dos posturas, que es independiente de la posición espacial y es independiente a la estatura del paciente, es mediante el cálculo del ángulo existente entre cada par de huesos adyacentes, tanto en el esqueleto del paciente como en el que representa la pose objetivo. Si los ángulos



Figura 4: Esqueleto Estandarizado

calculados coinciden, entonces la pose corresponde con la esperada. El ángulo entre cada par de huesos adyacentes es calculado mediante la Ley del Coseno, la cual dicta: que dada la distancia de dos catetos adyacentes en un triángulo y el ángulo formado entre ambos, se puede obtener la distancia del cateto opuesto aplicando la ecuación (1):

$$c^2 = a^2 + b^2 - 2ab \cos \alpha \quad (1)$$

Donde  $c$  representa el cateto opuesto,  $a$  y  $b$  los catetos adyacentes y  $\alpha$  el ángulo formado entre ambos. Dicha fórmula es una generalización del teorema de Pitágoras y puede ser reordenada como (2) para obtener el ángulo que se forma entre los catetos adyacentes.

$$\alpha = \arccos\left(\frac{a^2 + b^2 - c^2}{2ab}\right) \quad (2)$$

En el caso de uso actual, cada par de huesos conforman los catetos adyacentes de un triángulo imaginario compuesto por las tres articulaciones que forman los huesos. Entonces, el ángulo entre un par de huesos puede ser calculado con (2). Sin embargo, dos esqueletos con posturas distintas, pueden contener los mismos ángulos entre cada par de huesos, lo que acarrearía un falso positivo en la prueba de similitud. Por ende, el cálculo debe ser complementado con la evaluación de la relación espacial existente entre las articulaciones involucradas, tal como lo indica [19]. En la Figura 5, se puede observar como el ángulo del codo en ambas posturas es idéntico, sin embargo en la postura de la izquierda (A) la mano esta por encima del codo, mientras que en la imagen derecha (B) esta por debajo, lo que comprueba la importancia de evaluar las relaciones espaciales entre las articulaciones así como los ángulos.

<sup>2</sup><https://github.com/tensorflow/tfjs-models/tree/master/posenet>

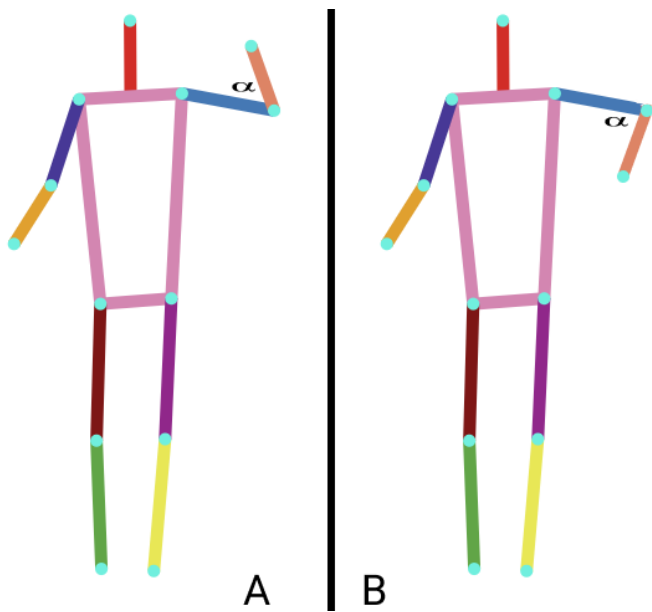


Figura 5: Dos Esqueletos con los Mismos Ángulos y Distintas Poses

El método que se propone en esta investigación es una modificación del discutido previamente. Dicho método evalúa la relación espacial entre cada articulación que compone el hueso y los ángulos entre cada uno, con un simple cálculo y un poco de preprocesamiento. El método se puede descomponer en los siguientes pasos:

- 1) Transformar la representación del esqueleto basada en puntos en el espacio, a una representación vectorial.
- 2) Calcular el ángulo existente entre los vectores que representen huesos homólogos en el esqueleto del paciente y el esqueleto objetivo.
- 3) Verificar que cada ángulo se encuentre dentro de un margen de tolerancia.

1) *Representación Vectorial:* En este primer paso, se toman las articulaciones de cada hueso del paciente y se genera el vector que compone cada uno de ellos. Es importante que el punto inicial y final de cada hueso en ambos esqueletos sea el mismo, debido a que dicha relación establece la dirección del vector. Esto quiere decir, que si se definió el vector que compone el antebrazo, como el vector que va del codo a la mano, en ambos esqueletos dicho hueso se debe calcular del codo a la mano y no en sentido arbitrario.

Cada vector calculado luego es normalizado. Al normalizarlo, se vuelve independiente de su posición inicial en el espacio y solo se conserva su orientación. La dirección del vector, encapsula la relación espacial entre cada articulación y permite validar que dos poses son idénticas.

2) *Cálculo de Ángulos:* En este caso, se calcula el ángulo existente entre la representación vectorial de los huesos homólogos en ambos esqueletos. Por ejemplo, al calcular el ángulo existente entre los vectores que corresponden al antebrazo del paciente y la pose objetivo, se compara si los mismos están orientados de la misma forma (ángulo entre los vectores igual a cero) y por ende corresponden parcialmente a la misma pose.

En este punto es que cobra vital importancia que los vectores que representan el mismo hueso, sean calculados utilizando el mismo orden en las articulaciones que lo comprenden, ya que en caso de estar alternados, correspondería con vectores cuyas direcciones son opuestas y el ángulo entre ellos no sería cero. Esta propiedad, permite discernir entre posturas cuyos ángulos internos son idénticos, pero la relación espacial entre las articulaciones que lo componen difiere.

El producto escalar es utilizado para calcular el ángulo entre ambos vectores. La ecuación (3) representa el producto escalar entre dos vectores, donde en el procedimiento actual no es necesario calcular la magnitud de los vectores ya que fueron normalizados en la etapa previa. En caso de que los vectores estén alineados, el ángulo entre ambos será cero y el producto escalar uno. El producto escalar es más eficiente que el método previo (2) ya que no requiere el cálculo del coseno y provee de la información necesaria para verificar la pose.

$$\cos(\alpha) = \frac{a \cdot b}{|a||b|} \quad (3)$$

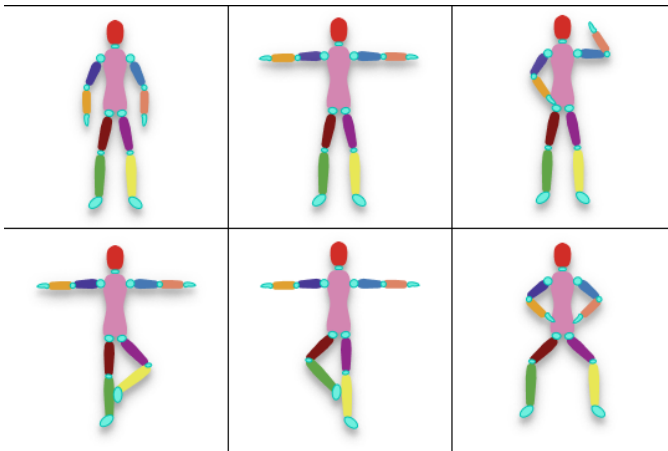
$$a \cdot b = a_x b_x + a_y b_y$$

3) *Verificación del Margen de Tolerancia:* Finalmente, se verifica si el producto escalar entre ambos vectores supera un margen de tolerancia impuesto por el personal médico. Dicho margen específica la diferencia en grados entre la dirección esperada del hueso y la lograda por el paciente. Al reducir el margen de tolerancia, se incrementa la dificultad del ejercicio.

### C. Maniquí Virtual

El personal médico dispone de un maniquí virtual para configurar la pose que el paciente deberá realizar en la terapia. El maniquí está compuesto por todos los huesos de la representación del esqueleto del paciente y el personal médico puede alterar la orientación o ángulo de cada uno de ellos, para generar una postura a seguir. Para facilitar el proceso de configuración, el esqueleto (del maniquí y del paciente) es representado como una estructura jerárquica (árbol). Cuando se aplica una transformación a un hueso, sus hijos se verán afectados por esta. Por ejemplo, si se rota el hueso correspondiente al brazo, el hueso del antebrazo rotará con este. El proceso de heredar las transformaciones del hueso padre es conocido como *Forward Kinematic* y se complementa con un conjunto de límites para los rangos de rotación de cada hueso, que previenen la configuración de poses anatómicamente incorrectas.

Finalmente, una pose inicial puede ser creada haciendo uso del algoritmo de detección de postura. El personal médico puede utilizar una foto de referencia o la cámara *web*, para que el sistema genere una pose inicial en base a la postura observada. Posteriormente, puede realizar los ajustes necesarios a la postura generada de manera manual. En la Figura 6 se puede observar el maniquí en distintas posturas que fueron configuradas a través del proceso descrito anteriormente.



**Figura 6:** Maniquí Provisto por la Aplicación para la Configuración de las Poses Objetivo

#### D. Aplicación del Médico

La configuración del maniquí virtual descrita anteriormente se realiza a través de una aplicación *web* orientada al personal médico.

Un ejemplo de la interfaz en uso se puede observar en la Figura 7. En la parte A se representa el proceso de configuración del ángulo del brazo. La interfaz consta de una barra de identificación en la parte superior (1), en la cual se encuentra el nombre de la aplicación, un identificador de nuevos eventos ocurridos en el sistema y el control de sesión del usuario. Un panel a la izquierda (2) con las opciones presentes en el sistema, como: panel de control, estadísticas, asignación de la sesiones de terapia, edición de ejercicios (seleccionada actualmente), manejo de pacientes, manejo de personal médico y vista de los dispositivos terapéuticos registrados en el sistema. En la parte superior central (3) se encuentra un campo de texto donde se identifica la postura creada, seguido de tres botones (4) que permiten filtrar los huesos del esqueleto que se despliegan en el área de trabajo, entre los huesos evaluados en la postura y los que no son evaluados. Posteriormente se encuentra el esqueleto con la postura actual (6), el médico puede interactuar con el esqueleto manipulando las articulaciones que definen un hueso o modificando el ángulo de apertura del mismo en la barra lateral derecha (12). Las articulaciones se presentan en color celeste, el hueso actualmente seleccionado en un color morado oscuro (7), los huesos activos en morado claro (8) y los inactivos en blanco. Para el hueso activo, se despliega de manera visual el ángulo que se esta evaluando, representado por un semicírculo purpura (10) y un cono que representa el margen de tolerancia asociado al hueso (9). En la esquina superior derecha se encuentra un indicador de orientación en el espacio (5) que especifica dónde está la izquierda, derecha, arriba y abajo del esqueleto actualmente visualizado. Dichas indicaciones cambiarán en base a la vista que se seleccione del maniquí, la cual puede ser modificada en la barra inferior central (11) entre la vista frontal y dos vistas laterales. Finalmente, el panel derecho contiene la configuración actual de todos los huesos del esqueleto, en el mismo el usuario podrá: activar y desactivar la evaluación de

un hueso para la postura actual, cambiar el ángulo esperado de un hueso y su margen de tolerancia.

La parte B ilustra cómo cambia el estado visual de la interfaz al seleccionar el antebrazo en el esqueleto (1). En este caso, el antebrazo pasa a ser de color purpura y el brazo cambia a un morado claro (2). Además, se despliega el ángulo evaluado para el antebrazo (4) y el cono que representa el margen de tolerancia (3).

#### E. Aplicación del Paciente

Finalmente, en la Figura 8 se puede observar un ejemplo de la interfaz del paciente en cada uno de los distintos estados que despliega a medida que el usuario realiza el ejercicio solicitado.

Como se puede observar, se utiliza el maniquí como representación visual de la postura que debe realizar el paciente. La posición del maniquí cambia en base a la posición del paciente dentro de la imagen, esto con la intención de que el maniquí no interrumpa la visibilidad del paciente al realizar el ejercicio. A su vez, el color del maniquí corresponde con los mismos colores desplegados en el esqueleto detectado sobre el paciente, esto procura facilitarle el proceso de identificación de las extremidades involucradas en la pose actual.

Se puede observar en la parte A de la Figura 8, como se despliega un cono con el margen de tolerancia asociado al ángulo esperado del hueso. El color del cono es el mismo que el del hueso al que corresponda para facilitar la identificación del movimiento a realizar. Cuando el usuario mueve la extremidad actual dentro del cono, el mismo desaparece y se despliega el cono correspondiente al margen de tolerancia de la siguiente extremidad mal posicionada (como se ilustra en la parte B y C). La representación jerárquica del esqueleto, permite el despliegue progresivo de los conos de tolerancia en el orden en el que el paciente debe realizar los movimientos para alcanzar la postura. Por ende, no se despliega el cono de tolerancia de un hueso hijo cuando el padre está mal ubicado. El despliegue progresivo de los elementos, disminuye la carga cognitiva que recibe el paciente al servir de guía en los pasos a seguir para alcanzar el objetivo. El resto de la secuencia muestra al paciente siguiendo las instrucciones que el sistema dispone de manera visual hasta que logra completar la postura y se le indica la conclusión del ejercicio mediante un mensaje.

Además de los conos que representan el margen de tolerancia, se despliegan un conjunto de flechas que indican la dirección general en la que el paciente debería mover la extremidad para alcanzar la posición deseada. Dichas flechas parten del punto medio de la extremidad actualmente evaluada (siguiendo la estructura jerárquica) y su dirección es calculada siguiendo la ecuación (4), que calcula el vector de desplazamiento escalado  $c$ , necesario para llevar la representación vectorial del hueso actual  $a$ , al esperado  $b$ .

$$c = (b - a)|a| \quad (4)$$

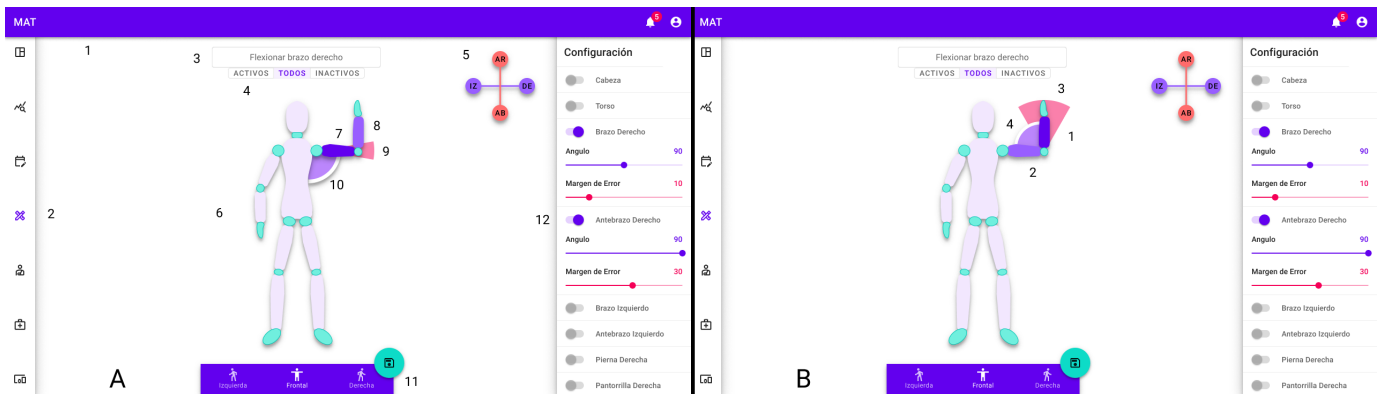


Figura 7: Vista de la Interfaz Utilizada por el Personal Médico para Configurar las Poses

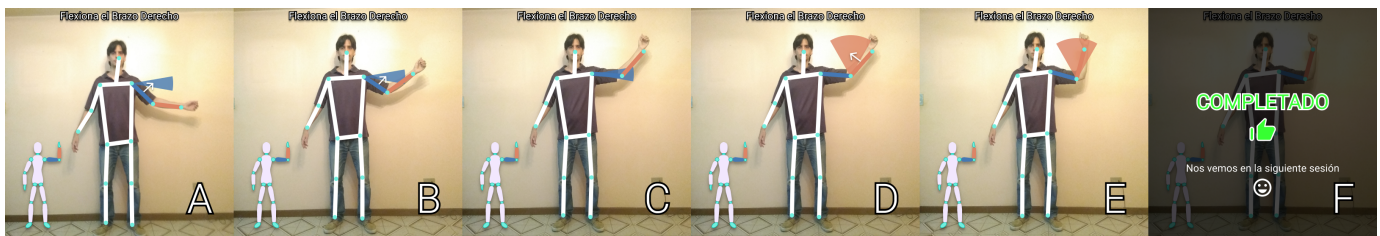


Figura 8: Secuencia de Interacción con la Aplicación del Paciente

### F. Integración con Plataforma IoT

La propuesta descrita hasta el momento, constituye un caso de uso específico de un dispositivo inteligente a integrar dentro de la arquitectura basada en *IoT* planteada en [4]. Para integrarlo dentro de dicha arquitectura es necesario: registrar el dispositivo, los tipos de ejercicios que soporta y sus parámetros, el tipo de dato correspondiente a los resultados que otorga el dispositivo y, tomar las consideraciones de comunicación necesarias para que el dispositivo pueda comunicarse efectivamente con los servicios de la arquitectura.

1) *Registro del Dispositivo*: El registro del dispositivo se realiza mediante una interfaz *web* diseñada para los desarrolladores de dispositivos inteligentes. En la Figura 9 se puede ver una fracción de dicha interfaz, en esta el desarrollador del nuevo dispositivo indica:

- El nombre del dispositivo.
- Una breve descripción sobre su función.
- Un conjunto de palabras claves utilizadas para buscarlo en la lista de dispositivos registrados.
- El sitio *web* del desarrollador (opcional).
- La lista de ejercicios que soporta.

Cada ejercicio en la lista de ejercicios soportados, contiene los siguientes datos:

- Un descriptor (nombre) de tipo de ejercicio.
- Descripción.
- Un conjunto de palabras claves para buscar ese tipo de ejercicios.
- Una lista con los parámetros de configuración requeridos.

Por cada parámetro a configurar se debe indicar:

- El tipo de dato del parámetro, entre los cuales se soportan los tipos: entero, flotante, lógico, cadena de caracteres,

Nombre

Manejador de Poses

Requerido

Descripción

Evalúa la postura del paciente con respecto a una pose objetivo

Requerido

Sitio Web

Opcional

Palabras Claves (#tags)

Postura ✕ Elasticidad ✕

Motricidad gruesa ✕ +

Ejercicios

Nombre

Figura 9: Configuración del Dispositivo

listas y objetos.

- Nombre del parámetro.
- Valor por defecto.
- Si el parámetro es obligatorio u opcional.

Dependiendo del tipo de dato del parámetro, se especifican restricciones a los valores que puede tener. Un tipo de dato numérico puede tener asociado un valor mínimo o máximo, una cadena de caracteres una longitud mínima o máxima y una lista tiene igualmente asociada una posible longitud mínima o máxima, además del tipo de dato por el que está compuesta. Construir un tipo dato “objeto”, consiste en indicar una lista con los atributos que contiene, cada uno de los atributos tiene la misma configuración que un parámetro tradicional, lo que permite construir un objeto cuyos atributos sean a su vez

objetos.

Finalmente, los datos que registra el dispositivo son configurados mediante una lista de parámetros. Cuando el dispositivo genere los resultados de la ejecución de un ejercicio, serán validados con la configuración especificada en este punto. Adicionalmente, el dispositivo indicará un porcentaje global de éxito asociado al ejercicio. Dicho porcentaje es calculado por el dispositivo, ya que el mismo contiene la lógica necesaria para evaluar la calidad de un ejercicio realizado. Este porcentaje será utilizado como un número referencial para evaluar la calidad de la ejecución del paciente.

De igual manera, posterior a la configuración del dispositivo, la aplicación utilizada por el personal médico tendrá la capacidad de crear un formulario con controles para cada uno de los parámetros que componen un ejercicio de la terapia (similar al panel 12 de la Figura 7). En el caso específico de las poses, los parámetros del ejercicio corresponden con una lista de los objetos que definen la dirección del vector que representa cada hueso del esqueleto. Debido a que los parámetros son de tipo espacial y tienen relaciones jerárquicas entre sí, resulta complejo configurarlos mediante campos numéricos tradicionales (configuración por defecto mostrada por la interfaz del médico). Por ende, se desarrolló una interfaz especial que facilite la configuración de dicho tipo de ejercicio, la cual se pudo observar en la Figura 7. En caso de que un nuevo dispositivo trabaje con parámetros complejos, que requieran el diseño de una interfaz particular para facilitar su manipulación, el desarrollador del mismo debe ponerse en contacto con el administrador de la arquitectura para realizar las adecuaciones pertinentes. Algunos ejemplos de datos complejos pudiesen ser: puntos geográficos (geolocalización), puntos en el espacio tridimensional o la representación de una postura mediante un maniquí virtual para el caso de uso particular del dispositivo presentado.

2) *Consideraciones de Comunicación:* Recordando [4], los dispositivos inteligentes se comunican con los servicios que componen la arquitectura, mediante el protocolo de comunicación *MQTT*<sup>3</sup> (*Message Queue Telemetry Transport* - Transporte de Telemetría de Cola de Mensajes).

El dispositivo enviará los resultados de un ejercicio a un tópico específico dispuesto para dicha función, suministrando el identificador único del ejercicio asignado y el resultado de dicho ejercicio. La estructura de un mensaje con estas características puede ser observada en la Figura 10, donde el campo “QUALITY” es el valor global de evaluación del ejercicio (rango [0 – 100]). Cuando la arquitectura registre los resultados, el dispositivo recibirá un mensaje indicando el nuevo estatus del ejercicio asignado como completado. De igual manera, si el dispositivo pierde la conexión con el servidor de mensajería, puede solicitar al momento de restablecer la conexión, las sesiones de terapia que pudieran haberle sido asignadas durante su tiempo de desconexión (ver Figura 11).

<sup>3</sup><http://mqtt.org>

<b>TOPIC:</b>	exercise/save-results
<b>ID:</b>	c187738b-819f-4c02-85f8-af926430e514
<b>RESULTS:</b>	steps: 1500
	time: 216.000
<b>QUALITY:</b>	98

**Figura 10:** Campos de un Mensaje que Registra el Resultado de un Ejercicio de un Dispositivo Cuenta Pasos

<b>TOPIC:</b>	exercise/get-assigned
<b>DEVICE ID:</b>	e96122a5-0573-4e70-ab66-976f405ceddc
<b>SINCE:</b>	1c6b1f91-30db-4c0c-82ff-da6107e26b45

**Figura 11:** Solicitud de Últimos Ejercicios Asignados

#### IV. PRUEBAS

Todas las pruebas que se aplicaron al sistema, se realizaron en un equipo con las siguientes características:

- Procesador: Intel(R) Core(TM) I5-8250U.
- Memoria RAM: 8 GB.
- Gráficos Integrados: Intel(R) UHD 620.
- Sistema Operativo: Windows 10.

Y un dispositivo *Android* con las siguientes características:

- Procesador: Snapdragon 430 1.4GHz.
- Memoria RAM: 3 GB.
- Sistema Operativo: Android 9.

##### A. Tiempo de Procesamiento

Se midieron los tiempos involucrados en cada una de las etapas que conllevan el proceso de identificar una postura, con la intención de detectar potenciales cuellos de botella que impidan que el producto final, pueda considerarse una aplicación en tiempo real. En la Tabla I se reflejan los tiempos de cada una de las etapas individuales del proceso de identificación de una postura. Como se puede observar, el tiempo total para procesar una postura es de 0,122 milisegundos, lo cual representa una tasa de 8.197 esqueletos procesados por segundo. Siendo la etapa más lenta de todo el algoritmo la normalización del esqueleto, que representa 43% del tiempo involucrado.

En la Tabla II se puede observar claramente las diferencias en velocidad de procesamiento de cada dispositivo de adquisición, donde se evidencia que el *Microsoft Kinect* es el método de adquisición más rápido, seguido por *PoseNet* y finalmente *OpenPose*. El tiempo de la identificación de la postura se mantiene constante ante todas las pruebas, debido a que solo depende del procesamiento de los esqueletos. El algoritmo de identificación de poses siempre realiza las mismas operaciones independientemente de la complejidad de la pose a evaluar, el procesamiento sería menor en caso de que la pose involucre

**Tabla I:** Tiempos por Etapas del Proceso de Comparación

Etapa	Tiempo	Fps
Esqueleto Estándar	0,042 ms	23.890
Normalización	0,053 ms	18.867
Comparación	0,027 ms	37.037
<b>Total</b>	0,122 ms	8.197



**Tabla II:** Tiempos de los Métodos de Adquisición

Método Adquisición	Detección Esqueleto	FPS
Kinect	33,05ms	30
Open Pose (CPU)	12.106,62 ms	0
Pose Net (MobiNet V1)	89,12 ms	11
Pose Net (ResNet 50)	109,63 ms	9
Pose Net (Android)	159,42 ms	6

una menor cantidad de huesos. Sin embargo, en las pruebas realizadas se consideró el peor caso posible, que sería que todos los huesos del esqueleto estén involucrados en la pose.

En base a los resultados expuestos en ambas tablas, se puede constatar que la velocidad del sistema reside en la del dispositivo de adquisición y que este representa un cuello de botella. Claramente la mejor alternativa es el *Microsoft Kinect* ya que permitiría que la aplicación se ejecute a 30 fotogramas por segundo, lo cual es una velocidad más que aceptable para una aplicación interactiva. Sin embargo, ya que dicha alternativa involucra adquirir un dispositivo de adquisición especializado, se considera que la mejor alternativa es el uso de *PoseNet*, que aunque es más lento, puede trabajar con una cámara *web* convencional. La aplicación con *PoseNet* se ejecuta a 9 fotogramas por segundo en su configuración más precisa y lenta y, a 11 fotogramas por segundo en la configuración menos precisa. Dicho rango se encuentra dentro del espectro en el que el humano percibe una secuencia de fotos, como un vídeo y no como imágenes separadas. Dada la velocidad de procesamiento, más la naturaleza de la aplicación, la cual involucra alcanzar una postura gradualmente sin necesidad de movimientos bruscos, permiten que *PoseNet* sea una alternativa viable para distribuir el sistema sin necesidad de requerir *hardware* especializado. De igual manera, se considera factible su configuración para dispositivos móviles, en la que el sistema fue capaz de operar a 6 fotogramas por segundo.

Finalmente, *OpenPose* no es capaz de alcanzar si quiera un fotograma por segundo con la configuración de *hardware* probada, por lo cual corresponde a la opción menos recomendada para la distribución del sistema. Para que *OpenPose* sea viable requiere de una tarjeta gráfica de alto poder de cómputo y por ende un mayor costo monetario. Sin embargo, es importante aclarar, que fue la alternativa que logró detectar poses complejas con mayor facilidad.

### B. Calidad de Detección

Con el objeto de verificar la efectividad del sistema para discernir entre poses correcta e incorrectamente realizadas, se aplicaron pruebas sobre un conjunto de datos provisto por la Universidad Politécnica de Cataluña [20]. Los datos están compuestos por imágenes de 12 actores realizando un total de 10 poses de distintas complejidades. Entre los actores se encuentran 10 personas de sexo masculino y 2 de sexo femenino de distintas alturas y contexturas. Es importante destacar, que las imágenes que componen el conjunto de datos no fueron capturadas haciendo uso del dispositivo terapéutico propuesto en este trabajo. Por ende, los autores de las imágenes califican la pose realizada por el actor en base a su similitud general con respecto a la pose objetivo, lo que trae como consecuencia

**Tabla III:** Imágenes por Pose e Individuo

Individuo	Pose				
	1	2	3	4	5
1	38	49	33	43	37
2	33	41	32	36	36
3	25	26	26	27	31
4	28	30	36	25	30
5	34	37	42	35	31
6	33	34	28	34	35
7	37	27	34	29	25
8	47	31	37	39	31
9	30	39	35	34	27
10	33	34	26	30	33
11	35	32	36	25	28
12	28	23	28	27	29

que dos posturas que no son realizadas exactamente igual, sean calificadas como la misma pose. La Figura 12 ilustra el problema descrito anteriormente, donde los 3 actores están realizando la misma pose; sin embargo, se puede observar claramente que existen diferencias entre las ejecuciones de cada uno.

Aunque el conjunto de datos está agrupado en 5 poses básicas y 5 poses complejas, en las pruebas se utilizó exclusivamente el grupo de poses básicas (ver Figura 13). Esta decisión se tomo en base a dos criterios: las poses calificadas como “básicas”, involucran más movimientos distintivos en las articulaciones de los sujetos; por otra parte, las poses “complejas” requieren un mayor esfuerzo de resistencia y equilibrio por parte de los actores (mantener piernas flexionadas o elevar una pierna en el aire), estas características no las hacen más complejas en cuanto a la evaluación de ángulos se refiere, pero si enfatiza la variabilidad en la ejecución de la pose entre cada sujeto, lo que acarrea mayores errores de calificación referentes a la exactitud de la pose realizada, por parte de los autores del conjunto de datos.

Para realizar las pruebas se seleccionó un individuo como referencia y por cada pose se eligió una imagen que identifique la misma, por lo tanto 5 imágenes (una por cada pose) fueron utilizadas como referencia para el individuo seleccionado. Una vez definidas las poses bases, se comparó cada pose contra todas las posturas realizadas por cada uno de los individuos en el conjunto de prueba, incluyendo el individuo utilizado como referencia. En la Tabla III se puede observar la cantidad de imágenes (posturas) grabadas por pose de cada individuo, en la cual se puede evidenciar que al realizar las pruebas, la proporción de poses incorrectas con la que se compara la pose referencial es muy superior a las poses correctas; por ejemplo, la pose 0 será comparada contra 1954 posturas, de las cuales 401 (20.52%) corresponden con el intento de realizar la misma pose (pose 0) y 1553 (79.48%) con el intento de realizar una pose distinta.

De las pruebas se extrajo la matriz de confusión presente en la Tabla IV. Para todas las poses se evaluó cada uno de los huesos del esqueleto que representa al usuario. La columna del margen de tolerancia, corresponde con los grados de diferencia tolerados entre cada par de huesos comparado en el esqueleto. Para todos los huesos se utilizó el mismo margen de tolerancia,



Figura 12: Errores Cometidos por los Usuarios para Imitar la Misma Pose Base



Figura 13: Poses Básicas Utilizadas en las Pruebas

Tabla IV: Matriz de Confusión

Margen de Tolerancia	VP	VN	FP	FN
$\approx 0$	5	7816	0	1949
5	173	7816	0	1781
10	278	7816	0	1676
15	478	7816	0	1476
20	750	7816	0	1204
25	1256	7816	0	698
30	1486	7816	0	468
35	1694	7789	27	260
40	1890	7745	71	64
45	1918	7613	203	36

sin embargo, en el dispositivo propuesto, el personal médico puede especificar un margen de tolerancia individual por cada hueso. Para calificar una pose como correctamente clasificada se utilizó las etiquetas (identificador de pose realizada) provistas en el conjunto de datos de prueba. Como se puede observar, con un margen de tolerancia muy pequeño ( $\approx 0$ ) solo se califican como correctamente realizadas 5 poses (VP), estas corresponden con la comparación que se realiza entre la pose seleccionada como referencia y ella misma, este comportamiento es esperado debido a que el grado de exactitud exigido por cada hueso es muy estricto, lo que hace que sea prácticamente imposible para un individuo realizar la postura exigida. A medida que se incrementa el margen de tolerancia, los errores cometidos al realizar la misma pose son permitidos y por ende la cantidad de verdaderos positivos incrementa y la cantidad de falsos negativos decrece. Por otra parte, se puede observar que al incrementar demasiado el margen de tolerancia aparecen falsos positivos, esto se debe a que la cantidad de grados de diferencia permitida entre cada hueso es tan laxa, que dos posturas diferentes pueden corresponder con el intento de realizar la pose seleccionada como referencia, este comportamiento es esperado y es deseado para evaluar pacientes que tengan rangos de movilidad muy reducidos.

Tomando en cuenta los motivos por los cuales surgen los falsos positivo y negativos, se puede concluir que el comportamiento es el esperado y el correcto, debido a que no representan una mala calificación por parte del algoritmo de comparación de poses, sino una mala asignación de la etiqueta de la pose en el conjunto de datos, afirmación que es posible ya que la etiqueta no corresponde con el margen de tolerancia asociada a la pose objetivo, ya que si el mismo es muy bajo, el nivel de exigencia para realizar correctamente la pose es muy alto y la mayoría de los intentos por realizarla serán rechazados y en caso contrario cualquier intento será calificado como correcto. Es importante destacar que aunque el conjunto de datos utilizado no esta diseñado con los niveles de exactitud requeridos para una aplicación médica, la cantidad de falsos positivos generados durante las pruebas realizadas es muy pequeña respecto a la muestra completa ( $\leq 10\%$ ). Se considera que para las características del sistema, un falso negativo carece de importancia respecto a un falso positivo, debido que un falso positivo representa aprobar la ejecución de la terapia del paciente aunque el mismo no intentase realizar la pose asignada. Sin embargo, un falso negativo involucra exigirle al paciente mayor esfuerzo para cumplir con el grado de exigencia impuesto por el personal médico. Como el dispositivo cuenta con diversos indicadores visuales del progreso del paciente: como el despliegue del esqueleto evaluado, el maniquí virtual de referencia, los conos de margen de tolerancia y flechas que indican la dirección del movimiento esperado y, que además dichos elementos son presentados de manera progresiva para guiar al paciente a cumplir con el objetivo, se considera que el sistema es capaz de ayudar al paciente a superar los falsos negativos.

Finalmente, es necesario indicar que durante las pruebas se utilizó el algoritmo de detección de esqueletos *PoseNet*, el cual presentó errores de detección en el 11.4% (223 imágenes) de las imágenes procesadas. En la Figura 14 se perciben

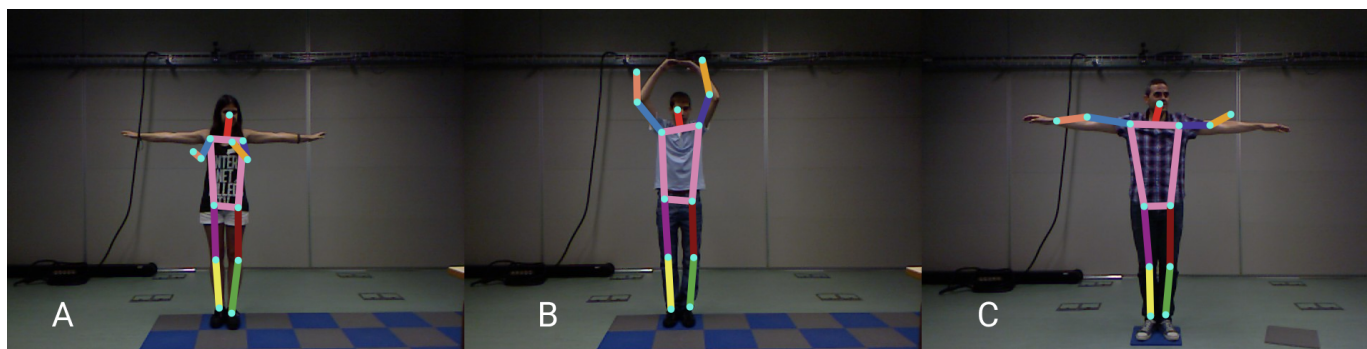


Figura 14: Errores de Detección de Esqueleto

algunos errores de detección, donde el esqueleto detectado no corresponde con la ubicación real del individuo. Estos errores de detección son producto de condiciones del ambiente en el que se realizó la prueba. Podemos observar que el tono de piel del individuo A puede confundirse con la pared de fondo; por otra parte, la disposición del cableado, tuberías y separadores en la pared de fondo pueden ser el motivo por el cual confunda la posición de alguna de las articulaciones de los individuos B y C. Ante estos problemas, se incluyó al sistema una lista de recomendaciones para minimizar los errores producto de una mala detección del esqueleto, como lo son:

- Utilizar ropa de distinto color al del fondo de la imagen.
- No utilizar ropa con rayas, preferiblemente prendas unicolor y no holgada.
- Procurar que el fondo de la imagen (lugar donde se realice la terapia) contenga la menor cantidad de elementos posibles.
- Remover elementos que obstruyan la visión del paciente.
- Contar con buena iluminación en el lugar donde se realizará la terapia.
- Alejarse de la cámara hasta que todo el rango de movimiento del paciente pueda ser capturado en la imagen.

Debido a que el paciente podrá observar el resultado de la detección del esqueleto sobre su cuerpo en tiempo real, se considera que será capaz de identificar una mala detección y los potenciales elementos que pueden estar generando dichos problemas en las imágenes capturadas por el dispositivo al consultar las sugerencias provistas.

## V. CONCLUSIÓN Y TRABAJOS FUTUROS

En el presente trabajo se describió y probó un dispositivo capaz de evaluar la postura del paciente y compararla contra una pose diseñada por el personal médico. El sistema propuesto es de bajo costo monetario y preciso. Además, está diseñado para disminuir la carga cognitiva generada en el paciente al suministrarle indicaciones graduales y claras de los pasos que debe seguir para poder realizar el ejercicio.

El dispositivo presentado se complementa con una interfaz web que le permite al personal médico configurar y adaptar los ejercicios propuestos al paciente, mediante el uso de un maniquí virtual que representa la postura objetivo para este.

Además, se indica el procedimiento para incluir un dispositivo dentro de la arquitectura basada en *IoT* previamente propuesta: desde su configuración, hasta las consideraciones que deben tomarse para realizar una comunicación efectiva con los servicios que comprenden la arquitectura.

Finalmente, se comprueba exitosamente la capacidad del algoritmo de comparación de poses diseñado, para evaluar correctamente las poses realizadas por el paciente. Siendo además capaz de trabajar con distintos mecanismos de adquisición de las características que definen su esqueleto.

Se propone como trabajo futuro, investigar más métodos de adquisición que mejoren el tiempo de respuesta y precisión del sistema, cumpliendo con la premisa de no involucrar la adquisición de *hardware* especializado costoso que dificulte su uso. De igual manera, se deben desarrollar otros dispositivos terapéuticos que sigan esta misma línea de pensamiento y fortalezcan la red de salud generada por la arquitectura *IoT*.

## REFERENCIAS

- [1] World Health Organization, *Rehabilitation*, World Health Organization, report, 2020.
- [2] K. Bandi, P. Pravati, P. Gopal, E. Jayasettiaseelon, Y. Sreekanth, M. Sridhar, and G. Gaurav, *Effect of Yoga Therapy on Heart Rate, Blood Pressure and Cardiac Autonomic Function in Heart Failure*, Journal of clinical and diagnostic research : JCDR, vol. 8, no. 1, pp. 14–6, 2014.
- [3] L. Bezants, *The Role of Exercise in the Management of Stroke*, 2016.
- [4] F. Moreno and W. Hernandez, *Sistema de Apoyo a la Terapia Física Basado en el Internet de las Cosas*, Revista Venezolana de Computación, vol. 6, no. 1, pp. 28–38, 2019.
- [5] A. Lotfi, C. Langensiepen, and S. Yahaya, *Socially Assistive Robotics: Robot Exercise Trainer for Older Adults*, Technologies, vol. 6, pp. 32, 2018.
- [6] Gov UK, *Exercises for Older People*, <https://www.nhs.uk/live-well/exercise/physical-activity-guidelines-older-adults/>, 2019.
- [7] I. Gregory and S. Tedjojuwono, *Implementation of Computer Vision in Detecting Human Poses*, in proceedings of the 2020 International Conference on Information Management and Technology (ICIMTech), pp. 271–276, 2020.
- [8] S. Yadav, A. Singh, A. Gupta, and J. Raheja, *Real-time Yoga Recognition Using Deep Learning*, Neural Computing and Applications, vol. 31, 2019.
- [9] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake, *Real-Time Human Pose Recognition in Parts from Single Depth Images*, in proceedings of Communications of the ACM, vol. 56, pp. 1297–1304, 2011.
- [10] Z. Cao, G. Hidalgo, T. Simon, S. Wei, and Y. Sheikh, *OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019.

- [11] Z. Cao, T. Simon, S. Wei, and Y. Sheikh, *Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields*, in proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017), Honolulu, HI, USA, 2017.
- [12] T. Simon, H. Joo, I. Matthews, and Y. Sheikh, *Hand Keypoint Detection in Single Images using Multiview Bootstrapping*, in proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017), Honolulu, HI, USA, 2017.
- [13] S. Wei, V. Ramakrishna, T. Kanade, and Y. Sheikh, *Convolutional Pose Machines*, in proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016), Las Vegas, NV, USA, 2016.
- [14] G. Papandreou, T. Zhu, N. Kanazawa, A. Toshev, J. Tompson, C. Bregler, and K. Murphy, *Towards Accurate Multi-person Pose Estimation in the Wild*, in proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017), Honolulu, HI, USA, 2017.
- [15] G. Papandreou, T. Zhu, L. Chen, S. Gidaris, J. Tompson, and K. Murphy, *PersonLab: Person Pose Estimation and Instance Segmentation with a Bottom-Up, Part-Based, Geometric Embedding Model*, in proceedings of the 2018 European Conference on Computer Vision (ECCV), Munich, Germany, 2018.
- [16] A. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications*, in proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017), Honolulu, HI, USA, 2017.
- [17] K. He, X. Zhang, S. Ren, and J. Sun, *Deep Residual Learning for Image Recognition*, in proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016), Las Vegas, NV, USA, 2016.
- [18] P. Green-Armytage, *A Colour Alphabet and the Limits of Colour Coding*, Color: Design & Creativity, vol. 5, pp. 1–23, 2010.
- [19] P. Borkar, M. Pulinthitha, and A. Pansare, *Match Pose – A System for Comparing Poses*, International Journal of Engineering Research & Technology (IJERT), vol. 8, no. 1, pp. 28–38, 2019.
- [20] J. Hidalgo and J. Casas, *Body Pose Dataset*, <https://imatge.upc.edu/web/resources/body-pose-dataset>, 2015.