

## Detección de caídas de adultos mayores en imágenes térmicas utilizando normalización angular, PCA y K-NN ponderado

### Fall detection of elder people on thermal imaging using angular normalization, PCA, and weighted K-NN

S. E. Ayala-Raggi <sup>a,\*</sup>, J. M. Roa-Escalante <sup>a</sup>, A. Barreto-Flores <sup>a</sup>, J. F. Portillo-Robledo <sup>a</sup>, L. G. Soid-Raggi <sup>b</sup>, V. E. Bautista-López <sup>c</sup>

<sup>a</sup>Facultad de Ciencias de la Electrónica, Benemérita Universidad Autónoma de Puebla, 72570, Puebla, Puebla, México.

<sup>b</sup>Facultad de Ciencias, Universidad Nacional Autónoma de México, Coyoacán, 04510, Ciudad de México.

<sup>c</sup>Facultad de Ciencias de la Computación, Benemérita Universidad Autónoma de Puebla, 72570, Puebla, Puebla, México.

#### Resumen

En este artículo se describe el diseño de un algoritmo capaz de clasificar imágenes térmicas de personas acostadas y no acostadas, con el objetivo de aplicarlo en un sistema de detección de caídas. Se diseñó un algoritmo de normalización automática en rotación, traslación y tamaño, aplicado a una base de datos de imágenes térmicas, con el propósito de obtener un nuevo conjunto de imágenes alineadas y así realizar una reducción de dimensionalidad mediante PCA. Se utilizaron secuencias de 100 fotogramas, y se produjeron tanto secuencias de caídas como de no caídas. Aplicando el clasificador K-NN ponderado para identificar la clase de cada fotograma, se obtuvo un vector de probabilidades de la clase acostado con 100 posiciones. Estos nuevos vectores se utilizaron como ejemplos de entrenamiento para un nuevo clasificador K-NN, el cual contiene ejemplos de vectores de probabilidad de caídas y no caídas. Aplicando la validación cruzada, nuestro sistema es capaz de reconocer caídas con un 91 % de precisión.

*Palabras Clave:* Reconocimiento de patrones, PCA, K-NN ponderado, Registro de imágenes, Detección de caídas

#### Abstract

In this paper we design an algorithm capable to classifying thermal images of people lying down and stand, to be applying it to a fall detection system. An automatic rotation, translation and size normalization algorithm was designed, to be applied of thermal image database, with the purpose of obtaining a new set of aligned images to performing a dimensionality reduction using PCA. Sequences of 100 frames were used, and both falling and non-falling sequences were produced. Applying the weighted K-NN classifier to identify the class of each frame, a probability vector of the lying class with 100 positions was obtained. These new vectors were used as training examples for a new K-NN classifier, which contains examples of falling and non-falling probability vectors. By applying cross-validation, the system is capable of recognizing falls with 91 % accuracy.

*Keywords:* pattern recognition, PCA, Weighted K-NN, Image registration, Fall detection

#### 1. Introducción

Las caídas son un evento fortuito que persiguen a los seres humanos durante su vida, dando pie a serias consecuencias que puede sufrir una persona si se encuentra en una edad avanzada. Según (OMS, 2021) las caídas se encuentran en el segundo lugar de mortalidad en el mundo, debido a esto, se han desarrollado varias técnicas para identificarlas, como el uso de sensores

térmicos, con los cuales se pueden aplicar técnicas de procesamiento de imágenes y aprendizaje máquina para identificar estos comportamientos.

El uso de sensores térmicos tiene ventajas importantes como una mejor segmentación de los objetos de interés, mientras éstos emitan radiación térmica. Además, se pueden evitar

\*Autor para correspondencia: [salvador.raggi@correo.buap.mx](mailto:salvador.raggi@correo.buap.mx)

**Correo electrónico:** [salvador.raggi@correo.buap.mx](mailto:salvador.raggi@correo.buap.mx) (Salvador E. Ayala-Raggi), [jesus.roae@alumno.buap.mx](mailto:jesus.roae@alumno.buap.mx) (Jesús Manuel Roa-Escalante), [aldrin.barreto@correo.buap.mx](mailto:aldrin.barreto@correo.buap.mx) (Aldrin Barreto-Flores), [francisco.portillo@correo.buap.mx](mailto:francisco.portillo@correo.buap.mx) (José Francisco Portillo Robledo), [lgsoidraggi@gmail.com](mailto:lgsoidraggi@gmail.com) (Lourdes Gabriela Soid Raggi), [veronica.bautistalo@correo.buap.mx](mailto:veronica.bautistalo@correo.buap.mx) (Verónica Edith Bautista López)

**Historial del manuscrito:** recibido el 21/05/2022, última versión-revisada recibida el 18/07/2022, aceptado el 09/08/2022, publicado el 05/10/2022

**DOI:** <https://doi.org/10.29057/icbi.v10iEspecial4.9344>



problemas que tienen las video cámaras como reflejos de luz y proteger de mejor manera la privacidad de los usuarios. En (Stojanovic et al., 2018) emplean imágenes térmicas y un algoritmo de seguimiento de objetos, destacando que cuando se utiliza tecnología de visión infrarroja, se espera que la mayoría de la radiación capturada provenga del objeto de observación, una característica de las imágenes térmicas es que no hay sombras y el ruido es completamente diferente al conocido en el espectro visible capturado por video cámaras. No se aprecian patrones de color, solo se detectan las variaciones de temperatura. Por lo tanto, las cámaras térmicas son muy importantes tanto en aplicaciones civiles como militares, especialmente en el dominio de la vigilancia y el monitoreo. Para utilizar un sensor térmico para monitoreo de caídas en una habitación, se debe contar con una cantidad extensa de ejemplos de entrenamiento si se quiere reconocer caídas utilizando métodos de aprendizaje automático, debido a que una caída puede darse en cualquier zona del área de captura, además, existe la posibilidad de que la caída suceda hacia cualquier dirección. Esta problemática nos conduce diseñar un algoritmo capaz de normalizar en rotación, tamaño y traslación del objeto de interés, para evitar el uso de una cantidad enorme de ejemplos de entrenamiento. Este método utiliza la obtención del centroide para detectar el objeto de mayor interés en la imagen, posteriormente se aplica una máscara circular eliminar los objetos que no se encuentren dentro de ella.

Dado que una caída está conformada por una secuencia específica de fotogramas, se propone dos etapas de clasificación utilizando K-NN Ponderado. La primera utiliza ejemplos de entrenamiento normalizados de dos clases: parado y acostado, de modo que puedan clasificarse los fotogramas de un conjunto de secuencias normalizadas y obtener el vector de probabilidad relacionado al comportamiento de acostado. Este conjunto consta de ejemplos equilibrados de caídas y de no caídas. La segunda etapa utiliza los vectores de probabilidad de las secuencias de caída y no caída como un nuevo conjunto de ejemplos de entrenamiento. Finalmente se realiza la clasificación utilizando secuencias de prueba y así determinar si la secuencia pertenece a una caída.

## 2. Trabajo relacionado

### 2.1. Registro de imágenes

(Jia et al., 2012) investigan el reconocimiento de huellas palmares, por lo que diseñaron un método que calcula el ancho de las palmas, logrando como principal aplicación la normalización de la escala de las huellas. Al binarizar la imagen de la mano y colocar un punto de referencia, utilizan un algoritmo de seguimiento de límites, donde basándose en la distancia entre la referencia y los bordes de la mano, encuentran los puntos mínimos para trazar una línea que una a dichos puntos, de esta manera normalizan la rotación de la mano. Para normalizar en escala utilizan los bordes inferiores de la mano para redimensionar el ancho de la palma a 300 píxeles, aplicando una detección de la línea de la vida mediante el uso de la transformada de Radon finita modificada. Finalmente realizan un recorte a la mitad de la palma para aplicar tres métodos diferentes de reconocimiento.

Algunos trabajos como en (Zulkifley et al., 2020), utiliza una base de datos de radiografías de manos para realizar una predicción de la edad ósea, usando la red Xception para entrenar un regresor con un conjunto nuevo de imágenes previamente segmentado y normalizado en rotación. La normalización de las radiografías de plantea en tres pasos, el primero realiza una segmentación de la región de la mano en las imágenes originales utilizando DeepLab V3 plus. El segundo paso, realiza la alineación del conjunto nuevo de imágenes segmentadas utilizando cuatro puntos clave de la imagen, obtenidos al aplicar un regresor separable de redes neuronales convolucionales MobileNet V1. Al trazar una recta entre los puntos clave del centro de la muñeca y la punta del dedo medio, se calcula el ángulo necesario para rotar la mano. Siguiendo el mismo enfoque de normalizar radiografías, en (Hrzic et al., 2021) proponen XAOM, el cual es un método de alineación y orientación de imágenes de radiografías de 21 regiones corporales diferentes. XAOM consta de dos etapas, la primera etapa del método, alinear imágenes de rayos X, experimentamos con los siguientes enfoques: transformada de Hough, algoritmo de detección de línea rápida y método de análisis de componentes principales. Para la segunda etapa, experimentan con las adaptaciones de varias topologías de redes neuronales convolucionales para predecir correctamente la orientación de la imagen las cuales son LeNet5, AlexNet, VGG16, VGG19 y ResNet50.

En (Zitova, 2019) definen el registro de imágenes como el proceso de superposición de imágenes tomadas en diferentes momentos con el fin de ser analizadas, utilizando diferentes puntos de vista o utilizando dispositivos de captura diferentes. El registro de imágenes es un paso crucial cuando se combinan varias fuentes de datos de imágenes, la detección de cambios y la restauración de imágenes multicanal. Utilizando este concepto, se muestra en (Uchida, 2013) donde se aplicaron diferentes técnicas de procesamiento de imágenes y reconocimiento de patrones para analizar bio-imágenes. Se abordan tareas como la transformación de niveles de gris, binarización, filtrado de imágenes, segmentación de imágenes, seguimiento visual de objetos, flujo óptico y el registro de imágenes. Al utilizar el registro de imágenes mediante técnicas como PCA, pueden conocer si existe alguna diferencia incluso después de la adaptación y encontrar la diferencia intrínseca, con el objetivo de comprender la deformación relativa de una imagen con respecto a la otra imagen.

En trabajos como (Yoruk et al., 2005) proponen una normalización en la traslación y rotación de imágenes de manos. El método consta de segmentar la región de las manos y utilizando la distancia radial del contorno de la mano con respecto a un punto de origen alrededor de la muñeca pueden identificar la posición de las puntas de los dedos y sus valles. Cada dedo se lleva a su orientación normalizada girándolo alrededor de la articulación metacarpiano-falange. La posición de las articulaciones que se encuentran debajo de la palma es inicializada con las posiciones de un ejemplo de mano que tiene ubicaciones conocidas. Al contar con estas posiciones se adaptan a la mano real utilizando una transformación de escala, basándose en el ancho de la palma y la longitud del dedo medio.

### 2.2. Detección de caídas

El uso de sensores infrarrojos de baja resolución para el seguimiento y detección de caídas ofrece un gran desempeño cuando se utilizan algoritmos de aprendizaje automático, tal como en (Chen and Ma, 2015) donde utilizan dos sensores para implementar un algoritmo de seguimiento basado en AOA y para la detección de caídas utilizan un clasificador K-NN obteniendo una precisión de 93 %. En el trabajo (Yu et al., 2020), desarrollaron un sistema de detección de caídas utilizando un sensor térmico de 80x64 píxeles, y una Raspberry Pi 3 el sistema es capaz de identificar una caída basándose en la altura del cuerpo humano, solucionando el problema de la variación de la altura en el campo de captura mediante un algoritmo de retroproyección, dando como resultado una correcta clasificación entre una caída de estar de pie, sentado, en cuclillas, caminar y acostarse. En (Rafferty et al., 2019) presentan una revisión de diferentes métodos para la detección de caídas usando imágenes térmicas, tales como el proceso lógico, el cual identifica las manchas en las escenas, suponiendo que se trata de una persona, se filtran los píxeles que están fuera del rango de radiación del cuerpo humano, la detección está basada en calcular el largo de dicha mancha. Otro proceso trata del análisis de escena, este método utiliza una Red Neural Convolutiva para clasificar las escenas y está entrenada para identificar caídas y no caídas. Por último, se implementa una metodología basada en análisis de escenas compuestas, la cual es un híbrido de las dos propuestas anteriores, de modo que se incorporan aspectos del movimiento para permitir que el aprendizaje automático identifique los parámetros relacionados con las caídas. Otras investigaciones orientadas a la detección de estos comportamientos en sitios específicos generan soluciones para diferenciar actividades normales de una persona al entrar a un baño y caída simuladas, logrando una tasa de discriminación del 97.8 %, destacando que las caídas se pueden detectar en menos de un segundo mientras que el sistema puede proteger la privacidad del sujeto (Kido et al., 2009).

### 3. Diseño del sistema de detección de caídas

El sistema completo mostrado en la figura 1 consta de 8 bloques principales mostrado en la figura, el primer bloque hace referencia a la secuencia de entrada previamente capturada para ser clasificada, en el siguiente bloque se realiza una normalización en rotación, traslación y escala de los fotogramas de la secuencia, con la finalidad de aplicar posteriormente una reducción de dimensionalidad utilizando PCA, el siguiente bloque contiene un clasificador que identifica cada fotograma utilizando un K-NN ponderado y empleando una base de datos de ejemplos de entrenamiento que contiene comportamientos de acostado y parado, los cuales fueron previamente normalizados y reducidos en dimensionalidad, el resultado de esta clasificación permite extraer un vector de las probabilidades de cada fotograma de que el sujeto se encuentre acostado, este vector es utilizado como entrada en el siguiente bloque, el cual es otro clasificador K-NN, pero ahora utiliza vectores de probabilidad pertenecientes a secuencias de caídas y no caídas como ejemplos de entrenamiento, además de estar de igual forma reducidos en dimensionalidad. Finalmente tenemos el bloque del re-

sultado, donde la predicción indica si la secuencia de entrada fue una caída o no caída.

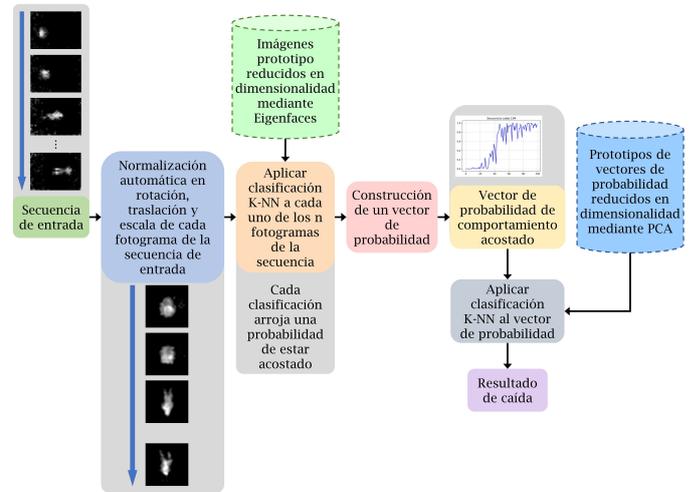


Figura 1: Diagrama de bloques sobre el flujo de datos propuesto en la construcción del sistema de detección de caídas utilizando secuencias, reducción de dimensionalidad y clasificación mediante K-NN.

### 3.1. Normalización automática en rotación, traslación y escala

La propuesta para normalizar nuestras imágenes se observa en la figura 2, muestra la secuencia de procesos que utiliza nuestro algoritmo, el tipo de imagen de entrada y la salida que se obtiene después de aplicar el algoritmo de normalización.

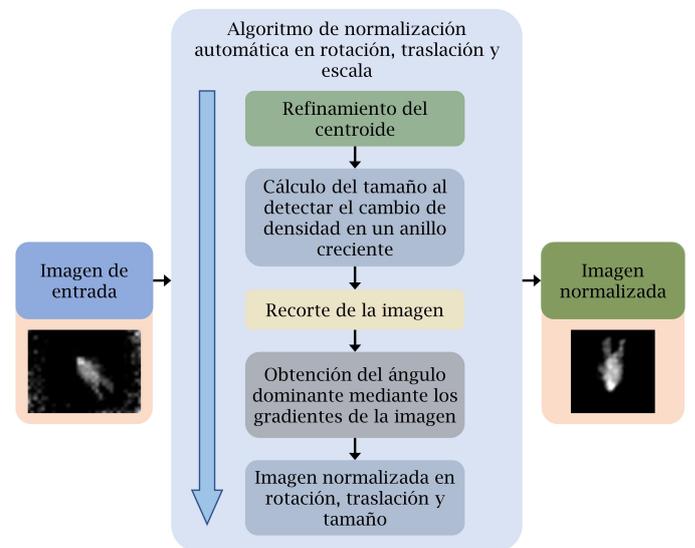


Figura 2: Diagrama de bloques del funcionamiento general del algoritmo propuesto para la normalización en rotación, traslación y escala de imágenes.

### 3.1.1. Normalización automática de traslación

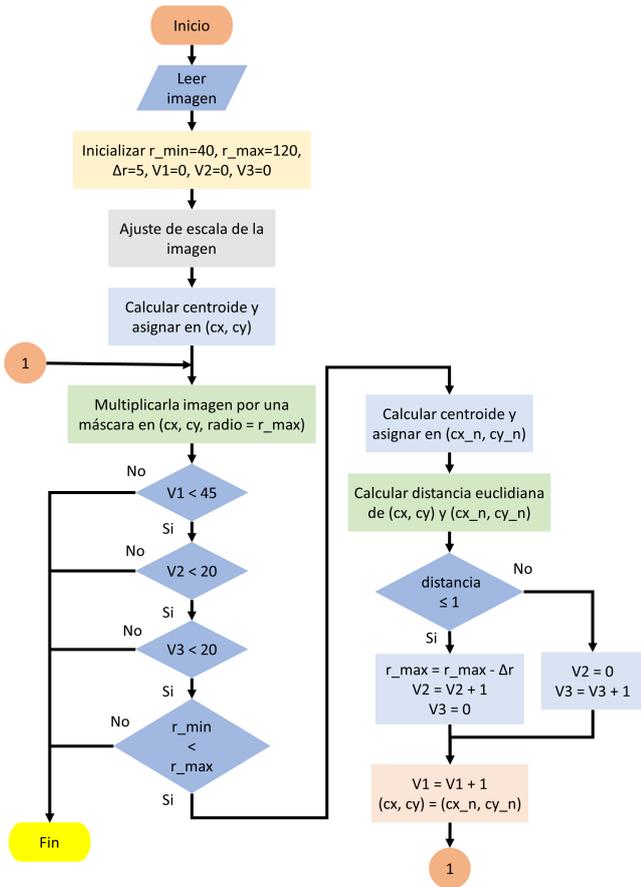


Figura 3: Diagrama de flujo del algoritmo propuesto para el refinamiento del centroide utilizando una máscara de radio  $r_{max}$  circular que permita eliminar el ruido fuera de ella.

El primer paso utilizado es el refinamiento del centroide, donde se realizan un redimensionamiento de la imagen, donde se espera que, si el conjunto de entrenamiento cuenta con imágenes en diferentes resoluciones, esto no afecte a la hora de emplear el algoritmo. Además, si se tuviera un cambio de sensor continuaría el buen funcionamiento del método desarrollado. Después, se procede a obtener el centro de masa de la imagen, debido a que se ha demostrado de forma experimental que el lugar con los píxeles de valor más alto de temperatura pertenece al cuerpo de la persona, dando lugar a que algunos ruidos térmicos en la muestra se terminen eliminando. Al ubicar el centro de masa, se procede a colocar una circunferencia de radio  $r_m$  en dicho punto, siendo por primera vez igual a 120 píxeles de radio, con el fin de volver cero todos los píxeles fuera de dicha circunferencia, posteriormente se realiza un ciclo donde el centroide es recalculado para la nueva imagen obtenida, utilizando tres variables que definen el criterio de paro para el algoritmo,  $V1$  que está relacionada a los ciclos máximos esperados del algoritmo completo,  $V2$  que define la cantidad máxima de ciclos en que la distancia euclidiana entre el centroide actual y anterior sea menor o igual a 1 para romper el ciclo,  $V3$  define la cantidad máxima de ciclos que pueden realizarse mientras se tenga que la distancia entre centroides sea mayor que 1. En cada ciclo transcurrido habrá un decremento del radio de la máscara circular de 5 píxeles, siendo el radio mínimo otro criterio

de paro, definido por un radio de no menos de 40 píxeles, el cual se conoce de manera experimental. Este proceso de refinamiento del centroide está descrito en el algoritmo mostrado en la figura 3. El proceso de actualización está mostrado en la figura 4, donde se muestra que al final del refinamiento se obtiene automáticamente el objeto de interés, además se muestra otra posible aplicación, la cual está basada en radiografías de manos, las cuales también es posible normalizar en traslación automáticamente, debido a que este tipo de imágenes también cuentan con un fondo negro y los objetos de interés están en escala de grises.

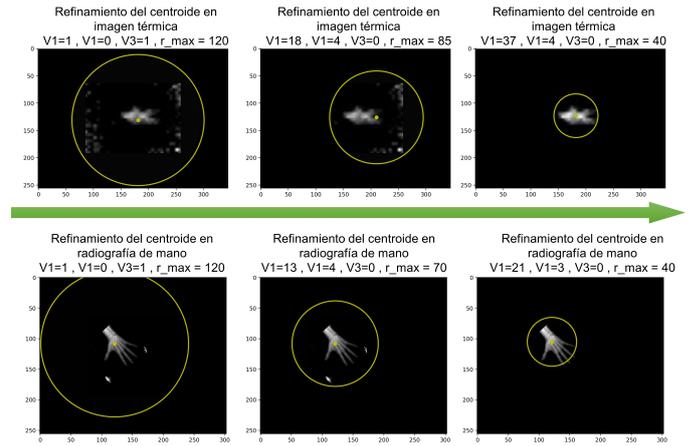


Figura 4: Proceso de búsqueda y ajuste del centroide comenzando de izquierda a derecha sobre la imagen para encontrar la región de interés en una caída.

### 3.1.2. Normalización automática de escala

Continuando con el proceso de normalización automática, ahora se realiza la obtención de la escala del objeto de interés, partiendo de la hipótesis de que la caída de densidad de los píxeles con forme se aleja del centroide es información suficiente para conocer el tamaño del objeto. Por este motivo se propuso medir la densidad en región de tipo anillo cuyo radio crece gradualmente. Así, el método comienza asignando a una constante  $c$  un valor de 1.32. Después se inicializa un conjunto  $R$  con valores para el radio de cada anillo, siendo en total 18 elementos, y el segundo conjunto contiene los valores de área de cada elemento del conjunto  $R$ . Para detectar la caída en el valor de densidad cuando los anillos crecen hasta llegar a los extremos del cuerpo humano, se proponen normalizar los valores de densidad  $D$  de los 18 anillos con el fin de homogeneizar la curva de disminución de densidad y así poder establecer un único umbral, por ello, primero se ordena el conjunto  $D$  comenzando por el valor mayor hasta el menor y tomamos los dos primeros valores máximos  $D_1, D_2$ , después se calcula el promedio  $\bar{m}$  entre ellos, y finalmente, normalizamos el conjunto  $D$  dividiendo cada elemento  $D_i$  por  $\bar{m}$  y así obtener el conjunto normalizado  $E$ . De forma experimental se tomó el valor de umbral de 0.2 para determinar la densidad mínima a la que debe llegar el algoritmo en todas las imágenes como se muestran en el Algoritmo 1.

El proceso del cálculo de la caída de densidad descrito en el algoritmo de normalización automática de escala está representado en la figura 5, la cual su radio crece con cada iteración, permitiendo encontrar la zona donde la densidad decrece hasta un umbral dado, asegurando que el objeto está contenido en ese radio descrito, además, se muestra el funcionamiento del

algoritmo aplicado a una radiografía de una mano, el cual también realiza la detección de la caída de densidad, debido a su similitud con un fondo negro y objetos en escala de grises.

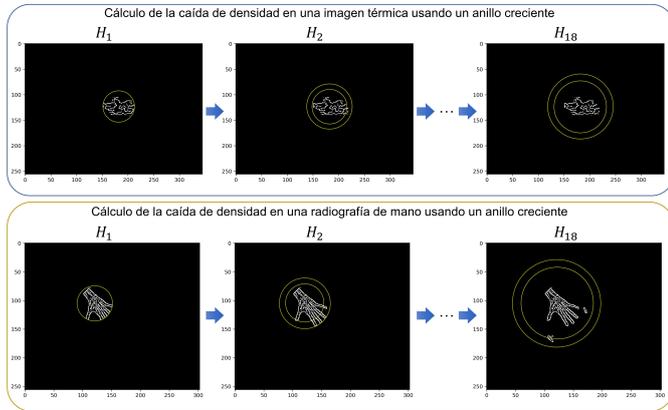


Figura 5: Proceso de cálculo de la caída de densidad automático utilizando un anillo creciente para normalizar la escala del objeto de interés en una imagen de caída. Se muestra el desempeño del algoritmo en una imagen térmica de una persona y en una radiografía de una mano.

**Algoritmo 1** Algoritmo para la normalización de la escala mediante el cálculo de la densidad en un anillo creciente.

**Entrada:** Imagen  $I(m, n)$ , bordes  $E(m, n)$

**Salida:** Imagen recortada  $I_E(m, m)$

- 1: **Inicializar:**  $c = 1,32$ ;  $R = \{r_1, r_2, r_3, \dots, r_{18} \mid r_i = 30c^i\}$
- 2: **Calcular:**  $A = \{A_1, A_2, A_3, \dots, A_{18} \mid A_i = \pi r_i^2\}$
- 3: **Crear máscaras:**  $M = \{M_1, M_2, M_3, \dots, M_{18}(m, n)\}$
- 4: **Calcular:**  $M_E = \{E(m, n)M_i(m, n) \mid M_i(m, n) \in M\}$
- 5: **Inicializar anillos y áreas:**  $H = \{H_1, H_2, H_3, \dots, H_{18}\}$   
 $H_A = \{H_{A1}, H_{A2}, H_{A3}, \dots, H_{A18}\}$
- 6: **Calcular la densidad:**  $D = \{D_1, D_2, D_3, \dots, D_{18} \mid D_i = \frac{H_i}{H_{A_i}}\}$
- 7: **Ordenar**  $D$  comenzando por el valor más grande
- 8: **Calcular** la media de los dos primeros valores de  $D$  como:  
 $\bar{m} = \frac{D_1 + D_2}{2}$
- 9: **Normalizar** las densidades de forma:  $E = \{E_1, E_2, E_3, \dots, E_{18} \mid E_i = \frac{E_i}{\bar{m}}\}$
- 10: **Hacer**  $j = 0$
- 11: **Mientras**  $x < 0,2$  **Hacer**
- 12:      $x = E_j$
- 13:      $j = j + 1$
- 14: **Fin Mientras**
- 15: **Seleccionar** circunferencia de radio  $r_j \in R$
- 16: **Calcular**  $I_E(m, n) = M(m, n)I(m, n)$  donde  $M(m, n)$  tiene radio  $r_j$

Una vez terminado el algoritmo 1, se obtiene la imagen normalizada automáticamente en traslación y escala mostrada en la figura 6. Como se observa, el objeto de interés se encuentra totalmente localizado y con un tamaño normalizado debido a que la caída de densidad es abrupta cuando el anillo creciente sale del objeto. cabe mencionar que las radiografías de manos lograron un desempeño similar con este algoritmo automático.

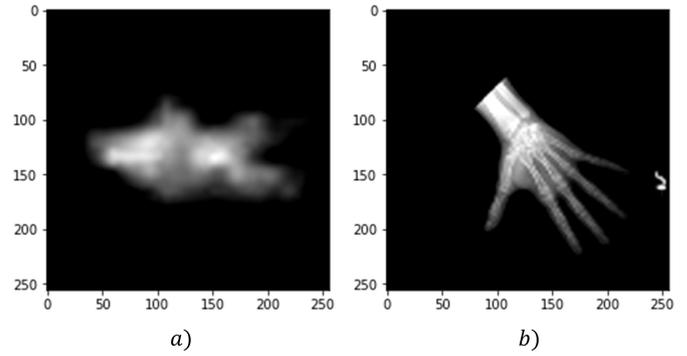


Figura 6: Recorte del objeto de interés utilizando el algoritmo de normalización automática de escala. En a) se observa el algoritmo aplicado a una imagen térmica y en b) a una radiografía de una mano

### 3.1.3. Normalización en rotación

Proponemos una técnica que permite corregir el ángulo de las imágenes de la base de datos. El primer paso es calcular el histograma de gradientes tanto horizontales como verticales de la imagen original  $I^g(m, n)$ . El mapa de gradientes  $G$  está compuesto por dos componentes, la primera define las magnitudes  $G_m(m, n)$  y la segunda contiene las direcciones  $G_a(m, n)$  de dichas magnitudes.

La obtención del histograma de gradientes orientados de la imagen recortada nos proporciona las direcciones en la que está distribuida la intensidad del gradiente, y se obtiene de manera vertical y horizontal para una imagen térmica de una persona y para una radiografía de una mano como se muestra en la figura 7.

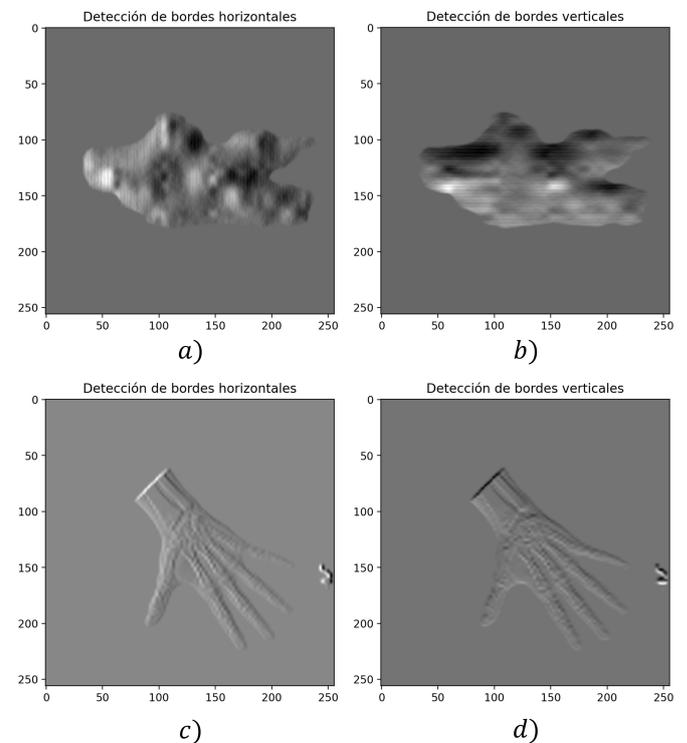


Figura 7: Obtención de los gradientes de una imagen de caída, a) gradientes horizontales, b) gradientes verticales.

Cuando utilizamos las componentes  $G_m$  y  $G_a$  es posible

calcular un histograma de ángulos con 360 intervalos, donde la componente en el eje ordenado se obtiene mediante la suma de todas las magnitudes del gradiente de los vectores dentro del intervalo angular correspondiente. Al mismo tiempo, cada columna de este histograma podría representarse como un vector cuya magnitud corresponde a la altura en el histograma y su ángulo corresponde a la etiqueta de su respectivo contenedor. Con esta información creamos un histograma polar de  $I^m(m, n)$ , que contiene un total de 360 vectores, en esta etapa es posible observar muchas variaciones en el seguimiento de las direcciones y magnitudes. Con la finalidad de suavizar este histograma original, creamos un nuevo vector  $H_i$  donde:  $i = 1, 2, 3, \dots, 360$ , donde en cada  $i$ -ésimo índice contiene la suma vectorial del conjunto de vectores contenidos dentro de una ventana angular que va del vector  $i$  al vector  $i + 90^\circ$  del histograma polar original. A partir de los nuevos vectores  $H_i$  se busca el vector de mayor magnitud  $h_{max1}$ .

La suavización del mapa polar original mostrado en la figura 8 b) y c) fue un método crucial para estabilizar la decisión del ángulo dominante, además permite agregar mayor resolución a las direcciones y magnitudes tanto de imágenes térmicas como de radiografías.

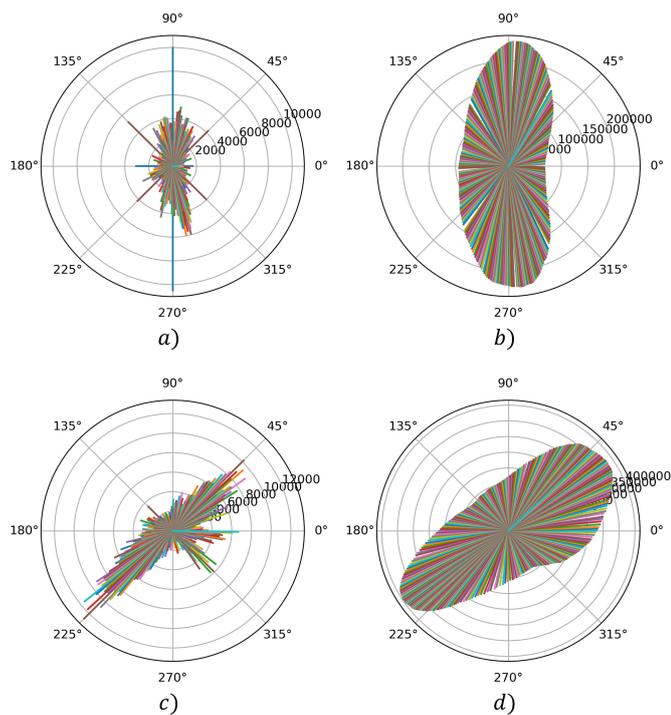


Figura 8: Histograma polar correspondiente a las direcciones y la sumatoria de magnitudes de las imágenes, a) histograma polar original para una imagen térmica, b) histograma polar suavizado para una imagen térmica, c) histograma polar original para una radiografía de mano, b) histograma polar suavizado para una radiografía de mano.

Después separamos la imagen original con una línea de dirección  $h_{max1}$  y tomando en la región opuesta al primer vector máximo. Ahora, tenemos que decidir cuál de estos dos vectores se considerará el principal vector dominante. Este enfoque permite centrarnos en la propiedad particular de las dos regiones separadas, a la primera mitad le denominamos  $C_1(m, n)$  y a la segunda  $C_2(m, n)$  mostradas en la figura . Se define una ventana deslizante  $P(m, n)$  de la mitad del tamaño de la imagen

$I^m(m, n)$ .

De esta manera, obtener el ángulo dominante de la imagen térmica se procede a dividirla en la dirección calculada como  $h_M$ , dando como resultado las dos imágenes  $C_1(m, n)$  y  $C_2(m, n)$  mostradas en la figura 9.

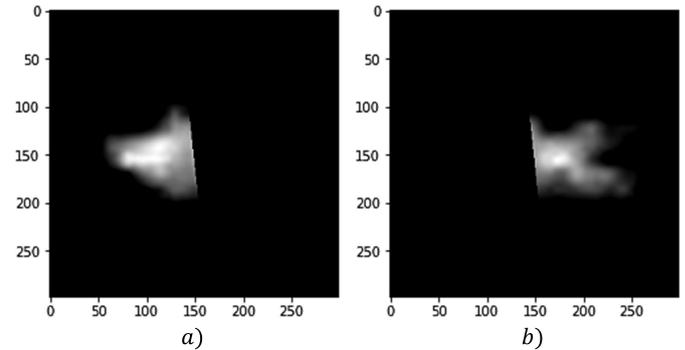


Figura 9: Imágenes resultantes de la separación de la imagen  $I^m(m, n)$  en dirección  $h_M$ . a) mitad denominada  $C_1(m, n)$ , b) mitad denominada  $C_2(m, n)$ .

Para todos los pixeles  $(m, n)$  de  $C_1(m, n)$ , centramos la ventana  $P(m, n)$  en  $(m, n)$  y calculamos la suma  $p$  de todos los valores de gris dentro de la ventana. Cada valor  $p$  se almacena en un vector  $S_1$ . Se realiza el mismo proceso para  $C_2(m, n)$  y se crea un vector similar  $S_2$ . Por otra parte, calculamos los valores medios  $\bar{p}_1$  y  $\bar{p}_2$  a partir de los vectores  $S_1$  y  $S_2$ . Finalmente, si  $\bar{p}_1 < \bar{p}_2$  entonces la dirección dominante será  $h_M = h_{max1}$  de lo contrario  $h_M = h_{max2}$ , de esta manera, se gira la imagen en dirección horaria los grados asignados a  $h_M$ , dando como resultado la imagen normaliza en rotación mostrada en la figura , el pseudocódigo de este proceso se describe en el algoritmo 2.

Por último, al ingresar al algoritmo de normalización automática en rotación una imagen térmica de una caída que contenga el objeto de interés en cualquier dirección o posición resultará en una imagen vertical, además, se muestra el desempeño obtenido utilizando radiografías de manos que se muestra en la figura 10, volviendo a nuestro algoritmo automático bastante viable para otras aplicaciones que utilicen imágenes con fondo negro y sea necesario detectar objetos de mayor tamaño que el resto de los objetos.

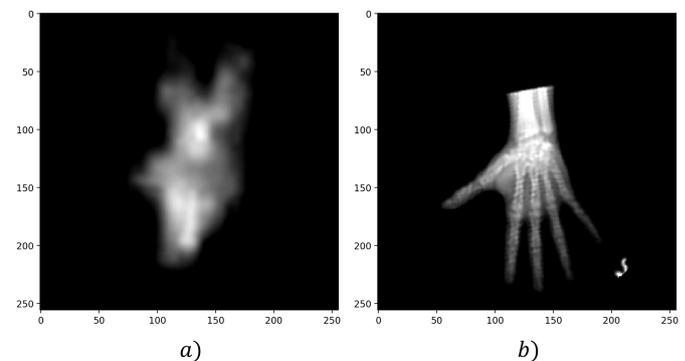


Figura 10: Imagen resultante de aplicar de manera conjunta los algoritmos de normalización en rotación, traslación y escala. a) imagen térmica normalizada, b) imagen normalizada de una radiografía de una mano

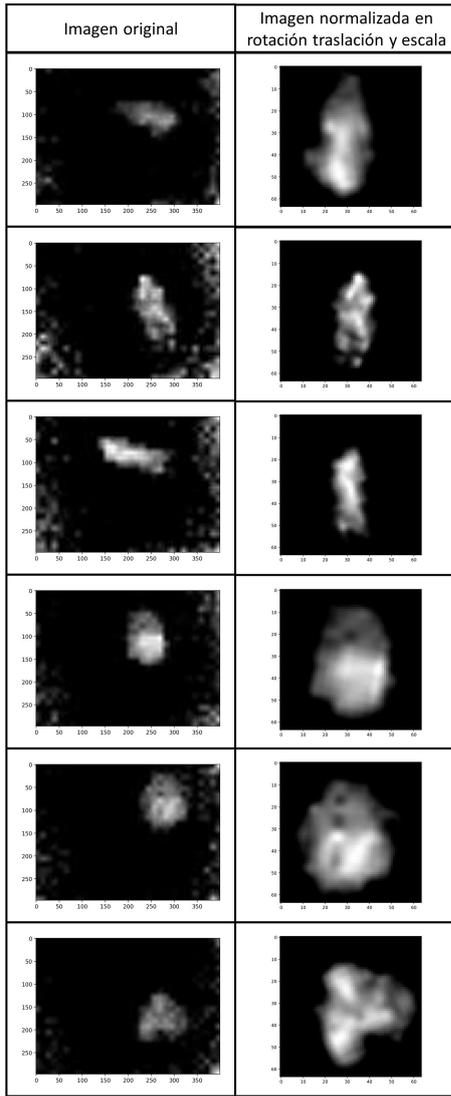


Figura 11: Resultado de algunos ejemplos a los que se les aplicó la normalización en rotación, traslación y escala, a la izquierda las imágenes originales y a la derecha el resultado normalizado.

En la figura 11, se puede observar el resultado en imágenes térmicas de utilizar en conjunto los algoritmos automáticos propuestos anteriormente, siendo una gran ventaja el contar con un fondo bastante diferenciable del objeto de interés, dadas las condiciones de usar un sensor térmico y que éste se encuentra también afectado por ruidos térmicos, es posible observar que la mayor parte del ruido se elimina gracias a la máscara circular que se utiliza en el algoritmo de normalización en traslación.

**Algoritmo 2** Algoritmo que permite identificar el ángulo dominante y determinar el ángulo de corrección mediante el histograma de gradientes orientados.

**Entrada:** Imagen  $I^g(m, n)$

**Salida:** Imagen corregida en ángulo  $I_1^S(m, n)$

1: **Hacer**  $I^m(m, n) = I^g(m, n)B(m, n)$

2: **Obtener** el mapa polar de  $I^m(m, n)$  en  $H$

3:  $h_{max1} \leftarrow \max(H)$

4:  $h_{max2} = h_{max1} + 180^\circ$

5: **Dividir**  $I^m(m, n)$  en dirección de  $h_{max}$

6: **Asignar** las dos partes de la imagen  $I^m(m, n)$  en  $C_1(m, n)$  y  $C_2(m, n)$

7: **Calcular** el valor medio  $\bar{g}$  de intensidad de gris en  $I^m(m, n)$

8:  $j = 1$

9: **Mientras**  $j \neq 2$  **Hacer**

10:     **Para cada** elemento  $m, n$  in  $C_j(m, n)$  **Hacer**

11:         **Si**  $C_j(m, n) > \bar{g}$  **Entonces**

12:             **Usar** ventana  $P(m, n)$  centrada en  $(m, n)$

13:              $p = \sum_{m,n} P(m, n)$

14:              $S_j \leftarrow p$

15:             **Fin Si**

16:         **Fin Para**

17:          $j = j + 1$

18: **Fin Mientras**

19: **Hacer**  $\bar{p}_1 = \text{mean}(S_1)$  y  $\bar{p}_2 = \text{mean}(S_2)$

20: **Si**  $\bar{p}_1 < \bar{p}_2$  **Entonces**

21:      $h_M = h_{max1}$

22:      $I_1^S(m, n) = \text{rotación horaria de } I^g(m, n) \text{ de ángulo } \angle h_M$

23: **Si no**

24:      $h_M = h_{max2}$

25:      $I_1^S(m, n) = \text{rotación horaria de } I^g(m, n) \text{ de ángulo } \angle h_M$

26: **Fin Si**

### 3.2. Clasificación mediante K-NN ponderado y reducción de dimensionalidad usando PCA

Para utilizar el clasificador K-NN ponderado se propone la reducción de dimensionalidad del conjunto de entrenamiento como se muestra en la figura 12, con el fin de obtener solo las características intrínsecas de los ejemplos de entrenamiento. Esta reducción está desarrollada mediante el método de PCA, en el cual las imágenes están representadas en una forma vectorial. Utilizando la imagen media resultante del conjunto de entrenamiento calculada en la ecuación 1 y la matriz de covarianzas vista en se pueden representar las imágenes originales mediante eigenvalores y eigenvectores utilizando la ecuación 2.

$$\Psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i \quad (1)$$

$$\Gamma_i - \Psi = \sum_{j=1}^k w_j u_j \quad (2)$$

De esta manera se muestra en la figura 12 la forma en que se propone normalizar y reducir en dimensionalidad el conjunto de prototipos de entrenamiento para realizar la clasificación de las secuencias de entrada tal como se muestra en el diagrama de bloques del sistema de detección de caídas.

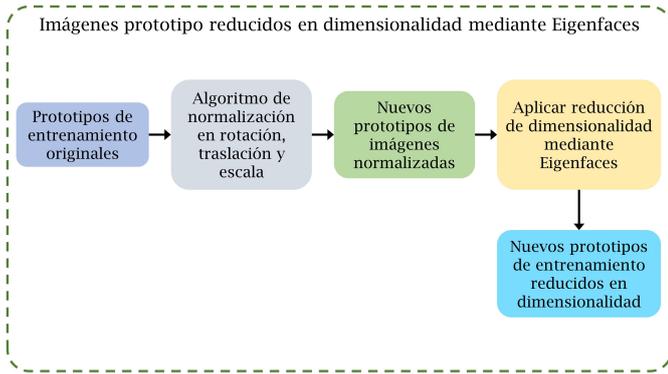


Figura 12: Diagrama de bloques del proceso de obtención de los nuevos prototipos de entrenamiento normalizados y reducidos en dimensionalidad de imágenes térmicas usando el método de Eigenfaces.

En la figura 13, se observa la propuesta de reducción de dimensionalidad aplicada a los vectores de probabilidad que se utilizan en la última clasificación utilizando K-NN ponderado, generando un procesamiento más rápido del algoritmo.

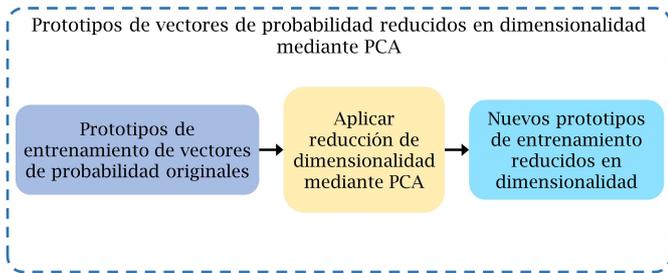


Figura 13: Diagrama de bloques del proceso de obtención de los nuevos prototipos de entrenamiento de vectores de probabilidad reducidos en dimensionalidad utilizando PCA.

Donde  $k$  representa la cantidad de eigenvalores utilizados de mayor a menor. Una vez realizada la reducción, es posible utilizar este nuevo conjunto de ejemplos de entrenamiento. El clasificador K-NN ponderado es un algoritmo de aprendizaje no paramétrico, esto quiere decir que cada predicción se hará tomando en cuenta todos los ejemplos de entrenamiento. Para hacer la clasificación, asumimos que todas las instancias corresponden a puntos en el espacio  $n$ -dimensional, si tomamos la instancia  $x$  y está descrita por el vector de características:  $(a_1(x), a_2(x), \dots, a_n(x))$ , donde  $a_r(x)$  denota el valor de la  $n$ -ésima característica de  $x$ , de esta manera, la distancia entre las instancias  $x_i$  y  $x_j$  está definida como  $d(x_i, x_j)$ , la cual corresponde a la distancia Euclidiana, que está definida como:

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^n a_r(x_i) - a_r(x_j)^2} \tag{3}$$

Para optimizar el valor de estas distancias se multiplicaron por un peso, el cual debe ponderar las distancias de modo que sean afectadas dando más peso a los vecinos que se encuentren más cercanos de la muestra y mientras la distancia crezca se reduzca el peso de los vecinos, utilizando la ecuación 4.

$$d = e^{-\frac{\text{distancias}^2}{2\sigma^2}} \tag{4}$$

Después se realiza la sumatoria de estas distancias de cada clase y se toma el argumento máximo. En esta primera etapa de clasificación se identifican dos clases de comportamiento, acostado y parado, de modo, que al obtener la probabilidad que se tiene de estar acostado es posible construir un vector de probabilidades, el cual contendrá la firma característica de una caída o una no caída.

#### 4. Resultados

Se obtuvieron un total de 4000 ejemplos del comportamiento acostado y 4000 ejemplos del comportamiento parado, con los cuales se realizó el entrenamiento del clasificador de la primera etapa. Durante la normalización en traslación obtuvo un buen desempeño a la hora de ejecutarse en las imágenes térmicas, ya que el algoritmo logra centrarse en el objeto de mayor tamaño y así eliminar parte del ruido de la imagen.

Los vectores de probabilidad obtenidos de la primera etapa de clasificación otorgan la información necesaria para determinar mediante un criterio simple si la secuencia es o no una caída. Este vector se forma utilizando la probabilidad de que la persona se encuentre acostada en cada fotograma de la secuencia, como se observa en la figura 14, el vector de probabilidad una secuencia de no caída y el vector de una secuencia de caída mostrado en la figura 15 son bastante diferenciables, permitiendo un buen desempeño del clasificador.



Figura 14: Vector de probabilidad de una secuencia de no caída.

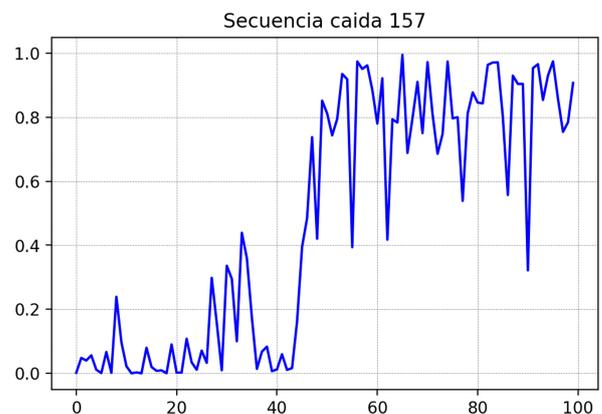


Figura 15: Vector de probabilidad de una secuencia de caída.

Al realizar la validación nuestro sistema de detección de caídas, se propuso realizar 100 pruebas donde los ejemplos de entrenamiento y test se eligieran siempre de manera aleatoria, obteniendo siempre una configuración de 352 secuencias de entrenamiento, las cuales están siempre equilibradas en 176 secuencias de caída y 176 secuencias de no caída, y un total de 60 secuencias de validación equilibradas de igual manera. En la figura 16, se puede observar el porcentaje de precisión de cada prueba realizada, siendo un 91.48 % la precisión media y la desviación estándar de un 3.04 %, utilizando un total de 11 vecinos cercanos, observando que a una cantidad pequeña de vecinos es posible obtener un mejor desempeño para identificar una caída, además, construimos la matriz de confusión, la cual se muestra en la figura 17, obteniendo una precisión de la predicción de caídas de un 0.906, una precisión de la predicción de no caídas de 0.923, una especificidad de 0.901 y sensibilidad de 0.927.

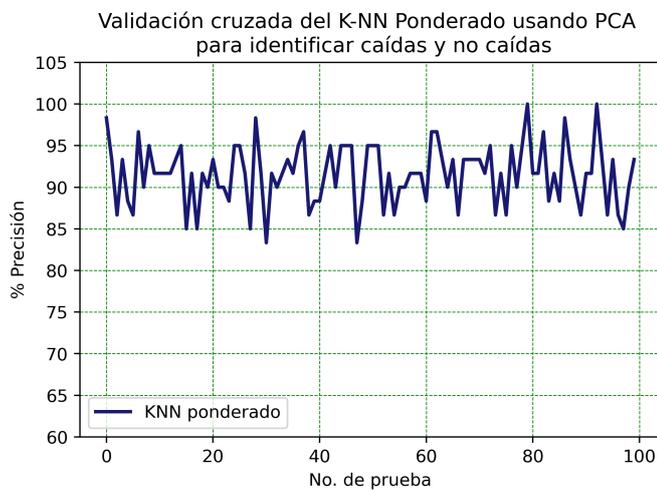


Figura 16: Resultados de la validación cruzada del clasificador K-NN ponderado con  $k = 11$ , 352 secuencias equilibradas (caídas y no caídas) de entrenamiento y un conjunto de validación equilibrado de 60 secuencias.

		Valores reales	
		Caída	No caída
Valores de predicción	Caída	47.12 %	4.86 %
	No caída	3.66 %	44.36 %

Figura 17: Matriz de confusión construida con los resultados de las 100 pruebas mediante validación cruzada.

## 5. Conclusiones

Nuestras principales contribuciones en este trabajo son 4, un algoritmo automático para la ubicación del centroide del objeto de interés en imágenes térmicas y radiografías de manos,

un algoritmo automático para normalizar la escala del objeto mediante la detección de la caída de densidad del objeto de interés con respecto al fondo negro en una imagen y que genera el recorte del objeto de interés, un algoritmo automático para la normalización angular del objeto de interés basado en el histograma de gradientes orientados y como última contribución está un sistema que utiliza los 3 algoritmos para detectar caídas utilizando vectores de probabilidad obtenidos de un clasificador K-NN y de la reducción de dimensionalidad. El algoritmo que se propone para normalizar en rotación, traslación y escala permite robustecer la clasificación de patrones, en este caso nos da la ventaja de buscar características o diferencias intrínsecas de cada clase mediante el registro de imágenes. La viabilidad en la clasificación de las secuencias utilizando técnicas como PCA y K-NN ponderado es alta, siendo estos resultados importantes para realizar un refinamiento del algoritmo. Además, conociendo que no se contó con una base de datos extensa, la normalización de los prototipos permitió prescindir de una gran cantidad de ejemplos de entrenamiento, ya que, al ejecutarse dicho algoritmo, se generan nuevos prototipos muy similares entre sí. Para el trabajo futuro se pretende mejorar el funcionamiento del algoritmo de normalización, con el fin robustecerlo y poder ser aplicado en otro tipo de imágenes que contengan un fondo negro, tales como imágenes médicas.

## Referencias

- Chen, W.-H. and Ma, H.-P. (2015). A fall detection system based on infrared array sensors with tracking capability for the elderly at home. In *2015 17th International Conference on E-health Networking, Application Services (HealthCom)*, pages 428–434.
- Hrzic, F., Tschauner, S., Sorantin, E., and ?tajduhar, I. (2021). Xaom: A method for automatic alignment and orientation of radiographs for computer-aided medical diagnosis. *Computers in Biology and Medicine*, 132:104300.
- Jia, W., Hu, R.-X., Gui, J., Zhao, Y., and Ren, X.-M. (2012). Palmprint recognition across different devices. *Sensors*, 12(6):7938–7964.
- Kido, S., Miyasaka, T., Tanaka, T., Shimizu, T., and Saga, T. (2009). Fall detection in toilet rooms using thermal imaging sensors. In *2009 IEEE/SICE International Symposium on System Integration (SII)*, pages 83–88.
- OMS (2021). Caídas.
- Rafferty, J., Medina-Quero, J., Quinn, S., Saunders, C., Ekerete, I., Nugent, C., Synnot, J., and Garcia-Constantino, M. (2019). Thermal vision based fall detection via logical and data driven processes. In *2019 IEEE International Conference on Big Data, Cloud Computing, Data Science Engineering (BCD)*, pages 35–40.
- Stojanovic, M., Vlahovic, N., Stankovic, M., and Stankovic, S. (2018). Object tracking in thermal imaging using kemelized correlation filters. In *2018 17th International Symposium INFOTEH-JAHORINA (INFOTEH)*, pages 1–6.
- Uchida, S. (2013). Image processing and recognition for biological images. *Development, growth differentiation*, 55(4).
- Yoruk, E., Dutagaci, H., and Sankur, B. (2005). Hand based biometry.
- Yu, L., Chen, H., He, H., Nie, H., Zhai, X., and Xiong, B. (2020). A fall detection system based on a thermopile imaging array and a back projection algorithm. In *2020 IEEE International Conference on Electro Information Technology (EIT)*, pages 060–065.
- Zitova, B. (2019). Mathematical approaches for medical image registration. In Narayan, R., editor, *Encyclopedia of Biomedical Engineering*, pages 21–32. Elsevier, Oxford.
- Zulkifley, M. A., Abdani, S. R., and Zulkifley, N. H. (2020). Automated bone age assessment with image registration using hand x-ray images. *Applied Sciences*, 10.