

## Navegación proactiva para asistencia tras caídas en un robot social

Sánchez-Girón, C.<sup>a,\*</sup>, Duque-Domingo, J.<sup>a</sup>, Gómez-García-Bermejo, J.<sup>a,b</sup>, Zalama, E.<sup>a,b</sup>

<sup>a</sup>Comité Español de Automática, Parc Tecnologic de Barcelona, Edifici U, C/ Llorens i Artigas, 4-6, 08028 Barcelona, España.

<sup>b</sup>Departamento de Automática, Ingeniería Electrónica e Informática, Universidad Politécnica de Madrid, C/ José Gutiérrez Abascal, nº2, 28006, Madrid, España.

<sup>c</sup>Departamento de Ingeniería de Ingeniería Eléctrica, Electrónica y Automática, Avda. de Elvas, s/n, 06006, Badajoz, España.

### Resumen

Este artículo presenta un sistema avanzado de detección de caídas y respuesta asistencial implementado en un robot social. Se propone una arquitectura híbrida que combina redes YOLO y ConvNeXt, junto con una estrategia de validación cruzada y votación ponderada, para mejorar la precisión de la detección en entornos domésticos. Para que el robot sea un elemento activo en el sistema, se ha integrado un módulo de acercamiento autónomo que permite al robot posicionarse frente a la persona caída tras una detección inicial, con el objetivo de realizar una nueva predicción y asistir directamente al usuario. Los experimentos realizados muestran que el acercamiento mejora significativamente la confianza del sistema en la clasificación y permite una interacción más efectiva. El sistema ha demostrado una buena precisión espacial y mejoras notables en la robustez de detección ante oclusiones. Este enfoque representa un avance hacia robots asistenciales proactivos, capaces de actuar en tiempo real ante situaciones críticas en el hogar.

*Palabras clave:* Visión por computadora, Integración de datos, Aprendizaje por conjuntos, Robots móviles, Detección de caídas.

### Proactive navigation for post-fall assistance in a social robot

#### Abstract

This paper presents an advanced fall detection and assistive response system implemented in a social robot. A hybrid architecture combining YOLO and ConvNeXt networks, together with a cross-validation and weighted voting strategy, is proposed to improve detection accuracy in home environments. To make the robot an active element in the system, an autonomous approach module has been integrated to allow the robot to position itself in front of the fallen person after an initial detection, with the aim of making a new prediction and directly assisting the user. The experiments conducted show that the approach significantly improves the system's confidence in classification and allows a more effective interaction. The system has demonstrated spatial accuracy and remarkable improvements in detection robustness to occlusions. This approach represents an advance towards proactive assistive robots, capable of acting in real time in critical situations in the home.

*Keywords:* Computer Vision, Data Integration, Ensemble Learning, Mobile robots, Fall detection.

## 1. Introducción

El número de personas mayores que eligen vivir en sus propios hogares, solas o con sus parejas, ha aumentado notablemente en los últimos años. Según el Instituto Nacional de Estadística (INE), en nuestro país, el número de personas mayores de 65 años que viven solas ha aumentado un 25.8 % en la última década (INE, 2023). Esto refleja un cambio en los patrones de vivienda, donde cada vez más personas mayores optan por mantener su independencia en sus hogares.

Es crucial asegurar que las personas mayores que viven so-

las estén en entornos seguros para evitar accidentes, especialmente caídas, que son la principal causa de lesiones en adultos mayores. Además, cuando alguien de la tercera edad se cae, el daño que sufre no es solo de forma física, sino que también se daña la confianza en su propia capacidad de estar solo. Por ello, es importante garantizar un entorno seguro para minimizar los riesgos de caídas y fomentar la autonomía de los mayores en sus hogares.

\*Autor para correspondencia: celia.sanchez-giron@uva.es

En este artículo se ofrece una propuesta innovadora para abordar este problema: un sistema de detección de caídas y navegación autónoma para asistir al usuario implementado en el robot social Temi (Temi, 2024). La integración de este sistema representa un gran avance hacia la creación de entornos domésticos inteligentes y seguros, especialmente dirigidos a mejorar la calidad de vida de personas mayores que viven solas. La investigación forma parte del proyecto EIAROB, que tiene como objetivo crear un ecosistema de Inteligencia Ambiental para apoyar el cuidado de larga duración en el hogar (EIAROB, 2022). El proyecto se articula en un sistema de inteligencia ambiental, apoyado en sensores, y soluciones robóticas para la atención a personas mayores, donde se integra la metodología que se va a presentar a continuación.

Este artículo se divide en diferentes secciones. En la Sección 2 se realiza una revisión bibliográfica sobre la tecnología actual de detección de caídas y robótica asistencial; en la Sección 3 se explica el procedimiento que se ha seguido para diseñar e implementar el sistema propuesto de detección de caídas y el sistema de aproximación autónoma al usuario; en la Sección 4 se muestran los resultados obtenidos tras realizar diferentes pruebas; por último, en la Sección 5 se discuten las conclusiones y la utilidad de esta propuesta.

## 2. Literatura previa

La seguridad de las personas mayores que viven solas ha impulsado el desarrollo de tecnologías para la detección de caídas, tradicionalmente basadas en sensores portátiles como acelerómetros, cuya eficacia depende del uso adecuado por parte del usuario (Sahni et al., 2025). Para superar estas limitaciones, se han propuesto soluciones sin contacto basadas en visión artificial, como redes neuronales convolucionales 3D (Fu, 2025) y modelos espacio-temporales (Shin et al., 2025), así como enfoques híbridos con sensores de profundidad (Xu et al., 2024).

Más allá de la detección pasiva, hay una tendencia a desarrollar robots sociales que integran capacidades de percepción y actuación. Este cambio se apoya en estudios como el de (Fischinger et al., 2016), que evaluaron la eficacia de robots como asistentes en el hogar, y (Coşar et al., 2020), que integraron visión artificial para el monitoreo continuo del usuario.

Un avance significativo es la incorporación de navegación autónoma en estos sistemas. Por ejemplo, algunos desarrollos recientes basados en las plataformas Pepper y Care-O-bot muestran cómo los robots pueden desplazarse hacia el usuario para verificar un evento detectado y ofrecer ayuda (Ribeiro et al., 2021).

(Oladele, 2025) también subraya que el futuro de la atención geriátrica está en la convergencia entre IA, monitoreo de salud y robótica social, resaltando cómo estas tecnologías pueden aumentar la calidad de vida en personas mayores.

A pesar de estos avances, la literatura muestra que aún existe una brecha crítica: muchos sistemas actuales son capaces de detectar eventos, pero no de actuar directamente sobre ellos. La integración de algoritmos de navegación autónoma con módulos de respuesta activa (como el acercamiento del robot tras una caída) representa una frontera tecnológica poco explorada, pero con un gran potencial.

Esta investigación aborda ese desafío mediante la integración de detección visual con un módulo de aproximación autónoma en el robot Temi, dentro del ecosistema EIAROB, para mejorar la verificación del evento y facilitar la asistencia directa al usuario.

Esta aproximación no solo busca aumentar la precisión del sistema mediante una segunda verificación visual, sino también facilitar la interacción directa con el usuario, para que el robot sea un elemento activo del sistema.

## 3. Metodología propuesta

En esta sección se describe el sistema propuesto para la detección de caídas en el robot social Temi, integrando la nueva funcionalidad de acercamiento al detectar a una persona caída.

### 3.1. Detección de caídas

Para la detección de caídas se utilizó una metodología novedosa que combina diferentes modelos neuronales junto con métodos de generalización para mejorar el comportamiento del sistema (Sánchez-Girón et al., 2024). El esquema del modelo de detección de caídas implementado se refleja en la Figura 2, donde se muestra la arquitectura completa.

La detección de caídas se realiza mediante la integración de las arquitecturas ConvNeXtSmall Backbone y YOLOv8s, utilizando métodos de generalización basados en el *validation voting*. El enfoque aprovecha las capacidades de YOLO para el reconocimiento de objetos en tiempo real y de ConvNeXt para el análisis de características profundas, tanto de la imagen como de los *bounding boxes* detectados.

El proceso comienza con la captura de imágenes por parte del robot, que se enviarán al sistema de detección de caídas para realizar la predicción. En primer lugar, un modelo YOLO identifica las personas y objetos de interés en la escena. Se han definido como elementos de interés sillas, camas y sofás, ya que se ha considerado que podrían influir en detectar la posición final de la persona. Para ayudar a la red a identificar si una persona se ha caído o está acostada en una cama o sofá, se ha incorporado una función que mide el nivel de superposición entre la persona y el objeto en la imagen, utilizando la técnica de Intersection Over Union (IoU). Con este método, se verifica si el área del rectángulo formado por el *bounding box* de la persona se cruza con el área del *bounding box* de cada objeto detectado.

$$IoU = \frac{Area_{persona} \cap Area_{objeto}}{Area_{persona} \cup Area_{objeto}} \quad (1)$$

Si el solapamiento es significativo (consideramos que es significativo cuando el índice IoU es mayor de 0.4), como el caso de la Figura 1, se mantienen las coordenadas de la persona en la imagen junto con los objetos que están solapados con ella, indicando una posible interacción. Si se detectan varios objetos del mismo tipo solapados en la imagen, conservaremos el solapamiento que tenga mayor área solapada porque consideramos que tendrá una probabilidad mayor de estar relacionado con la persona. En caso de que el objeto de interés no sea detectado o no se solape con la persona, el *bounding box* del objeto tendrá las coordenadas [0,0,0,0].

Así, tras la detección se obtiene una imagen recortada de la persona y cuatro conjuntos de coordenadas: una para el *bounding box* de la persona y tres para los *bounding boxes* de los objetos, presentes o no.

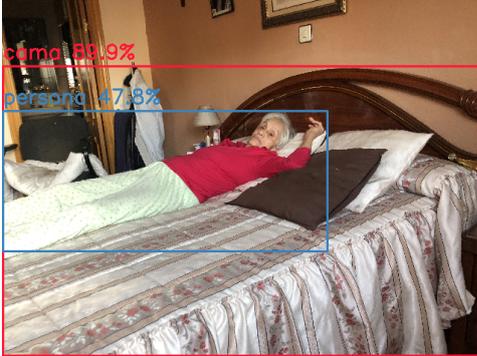


Figura 1: Ejemplo de solapamiento entre persona y cama en una imagen

Una vez que el modelo haya detectado las personas en la imagen y los objetos que están solapados con el área de las diferentes personas, las imágenes y las coordenadas obtenidas se introducirán en cada uno de los modelos generados mediante la técnica de generalización de *Cross Validation Voting* (CVV).

Esta técnica, como se expone en la investigación de (Domingo et al., 2022), permite mitigar problemas de sobreajuste y mejorar la capacidad de generalización del modelo. El método consiste en dividir el conjunto de entrenamiento en varios subgrupos o *slots*, en nuestro caso cinco, y luego entrenar un modelo para cada uno de estos subgrupos, aplicando la técnica de *early stopping* para evitar el sobreajuste. Una vez que se han entrenado varios modelos, se integran en un sistema de votación, *Soft Voting*, donde las predicciones de cada modelo entrenado se promedian, ponderando la clase con mayor probabilidad media (Sánchez-Girón et al., 2024). Esta combinación de modelos permite que el sistema se beneficie de las fortalezas de cada clasificador, compensando las debilidades individuales y logrando un rendimiento global superior.

Cada uno de los modelos generados por CVV presenta la misma estructura basada en una red ConvNext y una capa Fully Connected. El modelo ConvNext será interpretará la imagen de la persona recortada obtenida a partir de la red YOLO. La salida de este modelo será un vector de características de tamaño 512. Por otro lado, a la capa Fully Connected se le pasará como entrada los cuatro conjuntos de coordenadas correspondientes a los *bounding boxes* obtenidos anteriormente. Esta capa tendrá una entrada de tamaño 16, por estar constituida por el *bounding box* de la persona y los tres de los objetos (sofá, cama y silla) presentes o no. La salida de esta capa se concatenará con la salida del modelo ConvNext para introducir el vector resultante en una capa densa. Finalmente, esta capa devolverá un vector de dimensiones que coinciden con el número de clases, en nuestro caso, de tamaño uno ya que se trata de una salida binaria. Para entrenar los modelos se utilizó el conjunto de datos FPDS (Maldonado-Bascon et al., 2019), que contiene 3867 imágenes de personas caídas y 1005 imágenes de personas no caídas. Las imágenes fueron capturadas en entornos cerrados con posturas diferentes y oclusiones variadas.

### 3.2. Método de acercamiento

En este apartado se presenta la funcionalidad de acercarse a la persona caída detectada para ofrecer asistencia directa. El sistema de acercamiento consta de varias etapas, diseñadas para posicionar al robot de manera óptima. Para la navegación por el entorno, Temi usa un sistema de mapeo y localización (SLAM) propio del robot Temi con el que podemos acceder a las coordenadas locales del mapa definido.

El primer paso es calcular el ángulo que debe girar el robot para situarse de cara a la persona. Una vez que el robot detecta la caída, obtiene el *bounding box* de la persona en la imagen. Estas coordenadas se utilizarán para calcular el ángulo que debe girar el robot para situar a la persona caída en el centro de la imagen,  $\alpha$ . Así, teniendo en cuenta el ángulo de visión de la cámara del robot o *Field of View*, FOV, el ancho de la imagen,  $w$ , y la posición del centroide en  $x$  de la persona, podemos obtener el ángulo que debe girar el robot sobre sí mismo. En la Ecuación 2 se calcula el ángulo de giro a partir de las variables anteriores.

$$\alpha = \frac{\text{FOV}}{2} \left( \frac{\frac{w}{2} - X_{\text{centroide}}}{\frac{w}{2}} \right) \quad (2)$$

Este ajuste garantiza que el robot oriente su cámara hacia la persona antes de proceder al acercamiento.

El siguiente paso es calcular la distancia que hay entre el robot y la persona para que pueda aproximarse al punto exacto. Para calcular esta distancia se ha recurrido a técnicas de transformaciones homográficas para establecer relaciones de tamaño entre el mundo real y los píxeles de la imagen. Las transformaciones homográficas se aplican para representar un objeto visto desde un plano a otro diferente. De esta forma, puntos en una imagen podrán ser proyectados a un plano diferente mediante la matriz de transformación adecuada.

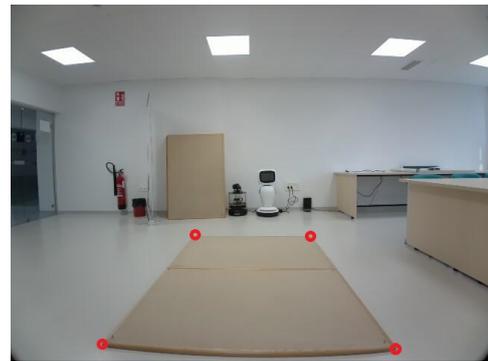


Figura 3: Imagen empleada para la transformación homográfica

En la Figura 3 se muestra la imagen que se ha utilizado para obtener la matriz de transformación homográfica. Para ello se empleó un tablero, con unas medidas conocidas, y se capturó una imagen del mismo con la cámara del robot a una distancia determinada. Se obtuvieron las coordenadas de las esquinas del tablero en píxeles, sistema de referencia de la imagen, y se asociaron con sus correspondientes coordenadas en el mundo real, tomando como punto de origen la posición de la cámara del robot en el momento de la captura. En la Figura 3 se ve claramente una distorsión en la parte inferior del tablero que podría corregirse para obtener una precisión mayor, del orden

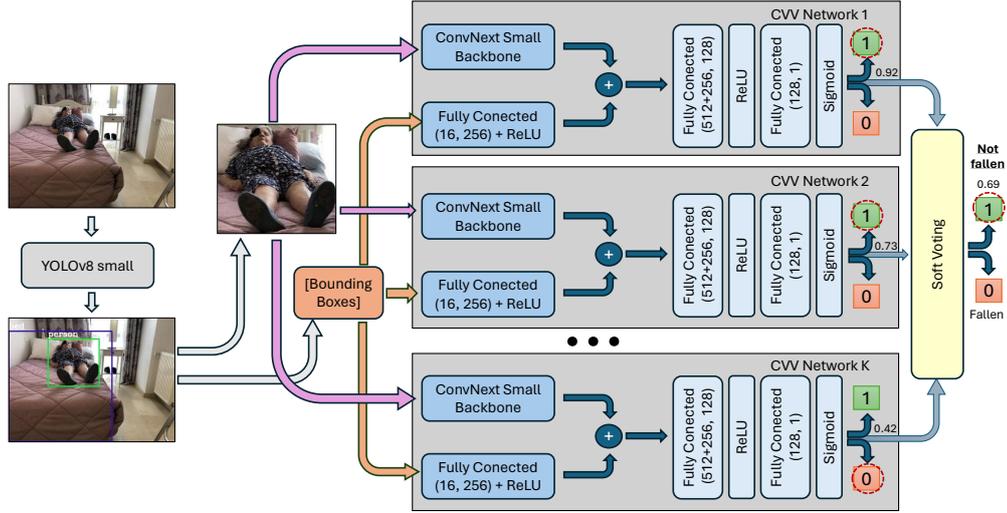


Figura 2: Esquema general del modelo de detección de caídas

de milímetros. Sin embargo, para el contexto de esta investigación se vio innecesario porque se ha comprobado que el posicionamiento del robot móvil se hace con una precisión del orden de centímetros y la distorsión de la imagen no influye en la correcta predicción del punto de destino. Con las coordenadas del tablero reales y en píxeles, se ha calculado la matriz de transformación homográfica empleando la función de la librería OpenCV de python `cv2.findHomography(srcPoints, dstPoints)`. Esta proporciona una matriz que transforma la perspectiva entre un plano de origen, píxeles de la imagen, a uno de destino, coordenadas en la vida real.

$$\mathbf{P}_{dst} = \mathbf{H} \cdot \mathbf{p}_{src} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} X'n \\ Y'n \\ n \end{bmatrix} \quad (3)$$

Una vez obtenida la matriz de transformación homográfica,  $\mathbf{H}$ , el sistema puede transformar cualquier punto de la imagen,  $p_{src}$ , a coordenadas del mundo real,  $p_{dst}$  como se muestra en la Ecuación 3. Hay que tener en cuenta que la homografía trabaja en coordenadas homogéneas y que el punto obtenido viene multiplicado por un factor de escala que no afecta al resultado final de la transformación, porque se cancela al dividir por este mismo factor.

Para calcular la distancia del robot a la persona, tomamos el punto más bajo del *bounding box* de la persona en la imagen, que estará en el plano del suelo, y lo convertimos a una coordenada del mundo real usando la matriz  $\mathbf{H}$  obtenida anteriormente. Finalmente, normalizaremos las coordenadas  $X'$  e  $Y'$  por el factor de escala  $n$  para obtener las coordenadas del mundo real como se muestra en la Ecuación 4.

$$X = \frac{X'}{n}, \quad Y = \frac{Y'}{n} \quad (4)$$

La distancia existente entre el robot y la persona en contacto con el suelo corresponde con la coordenada en el eje  $y$ .

Por último, calculamos la matriz de transformación *Robot-Base de carga* para transformar el punto que está en el sistema

de referencia local del robot, es decir, relativo a su posición y orientación actuales, al sistema de referencia base, el mapa del entorno. Esto es necesario porque el sistema de navegación del robot, y por lo tanto las órdenes de movimiento, está definido a partir del sistema de referencia base.

Para transformar un punto definido en el marco local del robot al marco de coordenadas global, aplicamos una matriz de transformación homogénea, Ecuación 5, que se basa en la posición del robot en el mapa,  $(x, y)$  y su orientación,  $\theta$ .

$$\mathbf{T} = \begin{bmatrix} \cos \theta & -\sin \theta & x \\ \sin \theta & \cos \theta & y \\ 0 & 0 & 1 \end{bmatrix} \quad (5)$$

Por último, obtenemos la posición en el sistema base  $(x_b, y_b)$  de un punto local  $(x_l, y_l)$  multiplicando el punto por la matriz en coordenadas homogéneas como se muestra en la Ecuación 6.

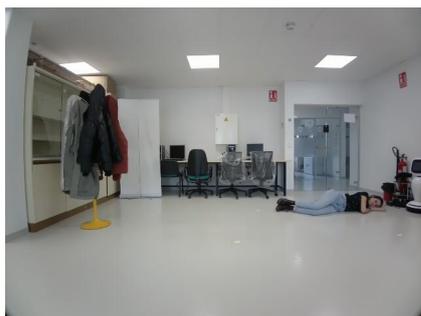
$$\mathbf{p}_{base} = \mathbf{T} \cdot \begin{bmatrix} x_l \\ y_l \\ 1 \end{bmatrix} = \begin{bmatrix} x_b \\ y_b \\ 1 \end{bmatrix} \quad (6)$$

La orden de movimiento final del robot estará constituida por las coordenadas obtenidas respecto al sistema base (posición) y la suma de la orientación actual del robot al ángulo de giro calculado (orientación).

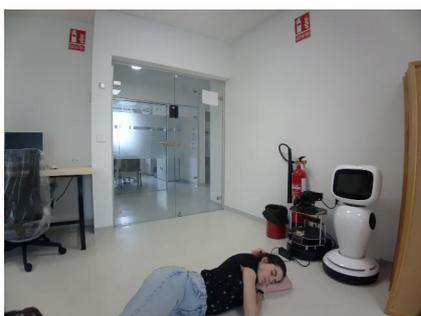
#### 4. Resultados

Para evaluar el comportamiento del sistema completo se han tenido en cuenta tanto los resultados obtenidos con el modelo de detección de caídas como la precisión del método de aproximación a la persona caída detectada. Para evaluar la eficacia del método de aproximación, se han realizado pruebas en las que la persona estaba tumbada en el suelo a diferentes distancias respecto al robot social. Las simulaciones se han repetido con diversas posiciones y distancias respecto a la posición inicial, en

concreto 2.5 metros y 4 metros, para analizar el método de acercamiento en situaciones de cercanía y lejanía. Se eligió 4 metros, ya que representa una separación adecuada para detectar una persona en el extremo opuesto de una habitación. Además, los ensayos realizados evidenciaron que el modelo YOLO comienza a presentar errores de detección cuando la persona se encuentra a más de 5 metros de la cámara del robot. Para evitar colisiones durante el acercamiento, se estableció una distancia mínima de 1.5 metros respecto a la posición estimada de la persona. En este contexto, la distancia de 2.5 metros se consideró apropiada para evaluar el comportamiento del sistema ante caídas cerca del robot.



(a) Antes de acercarse



(b) Después de acercarse

Figura 4: Comparación de imágenes según el método de acercamiento

Por un lado, se ha evaluado la precisión del sistema de aproximación a la hora de calcular y navegar hasta el punto deseado. En un entorno controlado, se establecieron varios puntos de referencia donde se situó una persona simulando una caída. El robot alcanzó la posición objetivo con un error medio inferior a 5 cm, lo que evidencia la precisión del sistema. En cuanto a la orientación, en todos los casos el robot se posicionó de frente al usuario, facilitando una interacción más natural y efectiva.

Por otro lado, se ha evaluado cómo el acercamiento del robot mejora la precisión en la detección de caídas. La Figura 4 muestra que al reducir la distancia y ajustar la orientación, el sistema obtiene una perspectiva más precisa, lo que favorece tanto la clasificación del estado del usuario.

En la Tabla 1 se recogen diferentes pruebas variando la distancia de detección y probando diferentes orientaciones de la persona respecto a la cámara. En las pruebas se han probado dos posiciones (paralelo a la cámara y perpendicular a la cámara) y tres ubicaciones (en línea con la cámara del robot, hacia la derecha y hacia la izquierda). En la Figura 4 por ejemplo, se ha representado una caída con posición paralela a la cámara, a 4 metros de distancia y ubicada a la derecha del centro de la ima-

gen. Esto ha dado lugar a 12 ensayos controlados para evaluar el acercamiento.

Para cada ensayo se ha calculado la confianza (probabilidad de que la clase corresponda a una caída) utilizando la función sigmoide, la cual retorna un valor entre 0 y 1. De este modo, se ha ponderado el valor de la confianza considerando que un valor próximo a 1 refleja una alta confianza del modelo en la predicción de caída, mientras que un valor cercano a 0 indica que el modelo está seguro de que no se trata de una caída.

El modelo de detección de caídas ha realizado una predicción acertada en prácticamente todos los casos, ya que todas las pruebas planteadas presentaban una persona caída; lo que demuestra la eficacia del modelo, independientemente del método de aproximación. Un patrón claro en la tabla es el aumento de confianza en la predicción *fall* cuando el robot se acerca, pasando de valores alrededor de 0.7 a más de 0.9 en algunos casos, lo que mejora la clasificación de la postura.

En ocasiones, los abrigos o ciertos objetos son reconocidos erróneamente como personas; son los casos cuarto y noveno de la Tabla 1 donde aparecen las predicciones *fall*, *notfall*, detectando a la persona real y a un abrigo. Aunque estas confusiones pueden persistir antes y después del acercamiento, suelen reducirse al sacar esos elementos problemáticos del campo de visión. Para mitigar los falsos positivos se podría fusionar la predicción con información de sensores, como sensores de presencia, que incluye el proyecto EIAROB (Ramos et al., 2024).



(a) Antes de acercarse



(b) Después de acercarse

Figura 5: Ejemplo de oclusión aplicando método de acercamiento

Por último, hay un caso en el que el sistema detectó al mismo individuo como dos caídas, el caso *fall*, *fall* de la Tabla 1. A pesar de ello, la predicción final sigue siendo acertada, dado que en ambos casos se registra el estado de caída, que es el estado que se prioriza.

Además, se ha demostrado que el sistema soluciona eficaz-

Tabla 1: Comparación del grado de confianza antes y después del acercamiento del robot.

Distancia	Antes		Después	
	Predicción	Confianza	Predicción	Confianza
4 m	fall	0.7943	fall	0.999
	fall	0.7187	fall	0.999
	fall	0.9965	fall	0.9954
	fall, notfall	0.557, 0.8118	fall	0.9886
	fall	0.9965	fall	0.9864
	fall	0.9886	fall	0.9996
2.5 m	fall	0.8967	fall	0.999
	fall, notfall	0.9569, 0.9965	fall, notfall	0.9871, 0.9986
	fall	0.6994	fall	0.9919
	fall	0.9957	fall	0.999
	fall	0.9101	fall	0.9781, 0.985
	fall	0.7789	fall, fall	0.9781, 0.937

mente problemas de oclusión, como se muestra en la Figura 5. En situaciones en las que la persona se cae detrás de un objeto, la vista inicial a distancia puede estar parcialmente bloqueada. Al acercarse, el robot permite una visualización completa del cuerpo y minimiza las falsas interpretaciones.

En resumen, los resultados muestran que la integración de un módulo de aproximación autónoma en el sistema de detección de caídas mejora la predicción del reconocimiento y resuelve eficazmente limitaciones como la oclusión o los falsos positivos.

## 5. Conclusiones

En este trabajo se ha presentado un sistema completo de detección y respuesta ante caídas, implementado en un robot social dotado de capacidades de visión artificial y navegación autónoma. A través de una arquitectura híbrida basada en redes YOLO y ConvNeXt, junto con técnicas de validación cruzada y *Soft Voting*, se ha logrado una detección robusta de caídas, incluso en condiciones de oclusión o escenas complejas.

Uno de los principales aportes del sistema es la funcionalidad de acercamiento inteligente tras la detección de una caída, que mejora la precisión de la predicción al obtener una nueva perspectiva más cercana, y facilita la interacción directa con el usuario. Este mecanismo se apoya en cálculos de orientación y distancia mediante homografía, integrándose de forma efectiva con el sistema de navegación del robot.

Los resultados experimentales han demostrado que el método de acercamiento no solo navega con la posición y orientación precisa para asistir a la persona caída, sino que también aumenta la confianza del sistema en sus predicciones, corrigiendo falsos positivos y mejorando la interpretación visual del entorno.

Este enfoque constituye un paso importante hacia entornos domiciliarios más seguros para personas mayores que viven solas, y abre nuevas líneas de trabajo en el desarrollo de una asistencia inteligente, proactiva y personalizada dentro del hogar.

## Agradecimientos

La investigación que se presenta en este trabajo ha recibido financiación del proyecto ROSOGAR PID2021-123020 OB-I00 financiado por MCI-NAEI/10.13039/501100011033/FEDER, UE, y del proyecto

EIAROB Financiado por la Consejería de Familia de la Junta de Castilla y León - Next Generation EU.

## Referencias

- Coşar, S., Fernandez-Carmona, M., Agrigoroaie, R., Pages, J., Ferland, F., Zhao, F., Yue, S., Bellotto, N., Tapus, A., 2020. Enrichme: Perception and interaction of an assistive robot for the elderly at home. *International Journal of Social Robotics* 12, 779–805.
- Domingo, J. D., Aparicio, R. M., Rodrigo, L. M. G., 2022. Cross validation voting for improving cnn classification in grocery products. *IEEE Access* 10, 20913–20925.
- EIAROB, 2022. Ecosistema de inteligencia ambiental para el apoyo a los cuidados de larga duración en el hogar mediante uso de robots sociales. <https://www.itap.uva.es/en/eiarob/>.
- Fischinger, D., Einramhof, P., Papoutsakis, K., Wohlkinger, W., Mayer, P., Panek, P., Hofmann, S., Koertner, T., Weiss, A., Argyros, A., et al., 2016. *Hobbit*, a care robot supporting independent living at home: First prototype and lessons learned. *Robotics and autonomous systems* 75, 60–78.
- Fu, F., 2025. A video-based fall detection using 3d sparse convolutional neural network in elderly care services. *Machine Graphics and Vision* 34 (1), 53–74.
- INE, 2023.
- Maldonado-Bascon, S., Iglesias-Iglesias, C., Martín-Martín, P., Lafuente-Arroyo, S., 2019. Fallen people detection capabilities using assistive robot. *Electronics* 8 (9), 915.
- Oladele, O. K., 2025. Ai in aging and elder care: Assistive technologies, health monitoring, and social robotics.
- Ramos, R. G., Duque-Domingo, J., Zalama, E., Gómez-García-Bermejo, J., 2024. Detección de actividades mediante modelos ocultos de markov jerárquicos. *Jornadas de Automática* (45).
- Ribeiro, T., Gonçalves, F., Garcia, I. S., Lopes, G., Ribeiro, A. F., 2021. *Charmie*: A collaborative healthcare and home service and assistant robot for elderly care. *Applied Sciences* 11 (16), 7248.
- Sahni, S., Jain, S., Saritha, S. K., 2025. A novel ensemble model for fall detection: leveraging cnn and bilstm with channel and temporal attention. *Automatika* 66 (2), 103–116.
- Shin, J., Miah, A. S. M., Egawa, R., Hirooka, K., Hasan, M. A. M., Tomioka, Y., Hwang, Y. S., 2025. Fall recognition using a three stream spatio temporal gcnn model with adaptive feature aggregation. *Scientific Reports* 15 (1), 10635.
- Sánchez-Girón, C., Zalama, E., Gómez-García-Bermejo, J. D.-D., 7 2024. Integración convnext-yolo mediante cvv para detectar caídas en robot social. *XLV Jornadas de Automática*. DOI: <https://doi.org/10.17979/ja-cea.2024.45.10788>
- Temí, R., 2024. URL: <https://www.robotemi.com/product/temi/>
- Xu, Z., Yang, X., Ruan, S., Dong, M., Sun, X., Qiu, W., Bi, S., 2024. Activity sensing system for elderly people living alone based on millimeter wave radar and depth camera. In: *2024 IEEE 14th International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*. pp. 468–473. DOI: 10.1109/CYBER63482.2024.10749183