

Seguimiento de movimiento de adultos mayores utilizando Transferencia de Aprendizaje Motion tracking of older adults using Transfer Learning

P. Bonifacio-Mariano a,*; E. Sánchez-Soto a^aDivisión de Estudios de Posgrado, Universidad Tecnológica de la Mixteca, 69004, Huajuapan de León, Oaxaca, México.

Resumen

La población mundial de adultos mayores se duplicará en el año 2050, lo que plantea desafíos en seguridad y calidad de vida, especialmente en entornos cerrados. En respuesta a esta necesidad, se han desarrollado sistemas de monitoreo enfocados en áreas como la identificación de patrones de comportamiento anormales, la prevención de caídas y la localización en interiores. Por este motivo, esta investigación presenta un sistema de visión por computadora para la detección automática de adultos mayores que involucra (a) un modelo preentrenado, el cual se especializa empleando la técnica de transferencia de aprendizaje en combinación con un conjunto de datos personalizado, y (b) se implementa un algoritmo de estimación de primer orden para la predicción del recorrido de la persona dentro del campo de visión a lo largo del tiempo. El enfoque en este trabajo contribuye a las soluciones tecnológicas que tienen como objetivo aportar al bienestar de los adultos mayores para mejorar su seguridad y calidad de vida.

Palabras Clave:

Transferencia de Aprendizaje, Visión por Computadora, Seguimiento de objetos, Seguimiento de trayectorias.

Abstract

The global population of older adults will double by 2050, posing challenges in terms of safety and quality of life, especially in indoor environments. In response to this need, monitoring systems have been developed that focus on areas such as identifying abnormal behavior patterns, preventing falls, and indoor localization. For this reason, this research presents a computer vision system for the automatic detection of older adults, which involves (a) a pre-trained model, further specialized using transfer learning with a customized dataset, and (b) a first-order trajectory estimation algorithm that predicts the person's path within the field of view over time. This approach contributes to technological solutions aimed at improving the well-being of older adults by enhancing their safety and quality of life.

Keywords:

Transfer learning, Computer vision, Object tracking, Trajectory tracking.

1. Introducción

De acuerdo con informes de la Organización Mundial de la Salud (OMS), se estima que en el año 2050, la población mundial de adultos mayores experimentará un incremento significativo, duplicándose respecto a los niveles actuales (OMS, 2015; 2024). En el presente, el crecimiento sostenido de este grupo etario plantea retos cruciales en referencia a su cuidado (monitoreo de la salud, asistencia médica y promoción de la autonomía). Ante esta realidad, surge la necesidad de desarrollar herramientas tecnológicas que permitan garantizar un en-

torno seguro, facilitando su autonomía y evitando riesgos físicos (Vélez *et al.*, 2024).

A través del uso de técnicas avanzadas para el seguimiento de personas, se han logrado avances considerables en aplicaciones como la videovigilancia (Nayak *et al.*, 2020; Deepthi, 2025). No obstante, la mayoría de las investigaciones se han enfocado en entornos abiertos, como calles y pasillos. Sin embargo, el cuidado de los adultos mayores en espacios cerrados (dormitorios, cocinas, baños y salas de estar) es fundamental, ya que estos son los lugares donde suelen pasar la mayor parte del tiempo (Chifu *et al.*, 2022).

* Autor para correspondencia: bomarj90@gmail.com

Correo electrónico: bomarj90@gmail.com (Perla Bonifacio-Mariano), esanchez@mixteco.utm.mx (Eduardo Sánchez-Soto)

Historial del manuscrito: recibido el 11/04/2025, última versión-revisada recibida el 01/09/2025, aceptado el 30/09/2025, en línea (postprint) desde el 03/10/2025, publicado el 05/01/2026. **DOI:** <https://doi.org/10.29057/icbi.v13i26.14921>



En consecuencia, la capacidad de implementar sistemas precisos y eficientes en interiores adquiere una gran relevancia al mejorar su calidad de vida, incrementando su seguridad y facilitando una supervisión continua (Sawik *et al.*, 2023; Georgoulas *et al.*, 2014; Güttsler *et al.*, 2014; Vélez *et al.*, 2024).

Por ello, en esta investigación se presenta un sistema de visión por computadora para la detección automática de adultos mayores, que involucra un modelo preentrenado (YOLOv8s) el cual se especializa por medio de la técnica de transferencia de aprendizaje (*transfer learning*, TL) y un conjunto de datos propio enfocado en personas de la tercera edad. El modelo se entrena con imágenes reales de adultos mayores en entornos abiertos y cerrados realizando múltiples actividades que reflejan sus rutinas diarias. Una vez que se detecta a la persona mayor, se implementa un algoritmo de estimación de trayectoria para realizar la predicción del recorrido de la persona dentro del campo de visión a lo largo del tiempo. Así, el sistema para el seguimiento de adultos mayores sienta las bases para su eventual integración en aplicaciones de mayor complejidad.

El contenido de este artículo se estructura de la siguiente manera: en la sección 2 se revisan trabajos relacionados con el tema central de esta investigación y se analiza cómo ha sido abordada esta problemática por distintos autores. En la sección 3 se detallan los pasos para la obtención del conjunto de datos AM-2025, desde la recopilación de las imágenes, las anotaciones y el preprocesamiento de las mismas; así como la aplicación de la transferencia de aprendizaje y el diseño del algoritmo de seguimiento. En la sección 4 se presentan los resultados obtenidos por el modelo especializado en la detección de adultos mayores y las pruebas realizadas con el algoritmo de estimación de la trayectoria. Finalmente, en la sección 5 se realiza la interpretación de los resultados presentados y se abordan las conclusiones de esta investigación.

2. Trabajos relacionados

2.1. Monitoreo de adultos mayores

El monitoreo de adultos mayores (*elderly monitoring*) abarca un conjunto de tecnologías y sistemas diseñados para apoyar a su supervisión, seguridad y calidad de vida. En (Kim *et al.*, 2022) se realiza una clasificación de estas tecnologías en seis funciones principales, las cuales se describen a continuación.

- Actividades diarias: monitoreo de actividades cotidianas tales como dormir o cocinar.
- Comportamientos anormales: identificación de actividades funcionales, psicológicas y sociales irregulares, tales como movimientos inesperados o patrones de comportamiento fuera de lo común.
- Deterioro cognitivo: detección de problemas cognitivos mediante el análisis de actividades diarias y patrones de comportamiento.
- Caídas: identificación de caídas mediante tecnologías basadas en visión (cámaras) y no basadas en visión (sensores de presión, radar).
- Localización de personas en interiores: seguimiento en tiempo real de la ubicación de los usuarios dentro del hogar.

- Calidad del sueño: evaluación y predicción de la calidad del sueño mediante datos de comportamiento.

A su vez, se identifican 16 tipos de tecnologías de sensores, destacando principalmente los de movimiento, de contacto y las cámaras de profundidad. Estos dispositivos se instalan en ubicaciones específicas del hogar, como techos, paredes, muebles y electrodomésticos, con el fin de recopilar datos sobre las actividades y el comportamiento de los adultos mayores.

Además, en (Nguyen *et al.*, 2024) se destaca la importancia de implementar sistemas no invasivos centrados en tecnologías como cámaras, señales Wi-Fi, sensores iniciales (IMU) y técnicas de fusión de datos para el seguimiento de actividades y localización de personas mayores, especialmente en hogares inteligentes o residencias asistidas.

En este ámbito, se han desarrollado soluciones para detectar patrones de actividad, caídas e inactividad prolongada. Por ejemplo, para monitorear la rutina diaria de personas mayores en (Fuentes *et al.*, 2021) se propone un sistema basado en dispositivos IoT y en (Ahmed y Abdulrazak, 2021) se presenta un sistema que emplea sensores térmicos, mientras que en (Zin *et al.*, 2021) se utiliza visión por computadora para detectar caídas o anomalías en la movilidad.

2.2. Localización en interiores

Un aspecto fundamental del monitoreo de personas mayores es la localización en interiores (*indoor localization*), este término ha sido objeto de análisis en diversos estudios, dado que permite identificar si la persona se encuentra en una zona de riesgo, como cerca de escaleras o espacios con condiciones peligrosas, y facilita la localización en situaciones de emergencia. En este sentido, en (Karkar *et al.*, 2021) se presentó Cam-Nav, un sistema de navegación en interiores basado en visión por computadora que utiliza características de patrones binarios locales multiescalares (MSLPB) para mejorar la precisión en el reconocimiento de lugares. En contraste, en (Chen, 2022) se usa tecnología Beacon para monitorear la seguridad de los adultos mayores en casa, con cámaras que detectan el uso de electrodomésticos de alto riesgo y alertan a los cuidadores. Por otro lado, en (Chen *et al.*, 2022) se propone un sistema de primeros auxilios que combina detección de caídas mediante visión y enrutamiento de rescate a partir de modelos de información de construcción (BIM). En (Lee *et al.*, 2024), se desarrolla un sistema que emplea etiquetas Bluetooth, sensores de techo y análisis de Big Data para rastrear en tiempo real los movimientos de residentes, personal y visitantes en residencias geriátricas. Finalmente, en (Mane *et al.*, 2025) se presenta un enfoque que integra segmentación de imágenes, extracción de características posturales, seguimiento del flujo de movimiento y algoritmos de clasificación para distinguir entre eventos críticos y situaciones no peligrosas para personas de la tercera edad.

2.3. Seguimiento de objetos/personas

Por otro lado, el seguimiento de objetos (*object tracking*) lleva el monitoreo un paso más allá. Esta tecnología, basada en cámaras inteligentes equipadas con visión por computadora y algoritmos de procesamiento de imágenes, permite seguir de manera continua y en tiempo real el movimiento de la persona a

medida que se desplaza por un espacio. A diferencia de la localización en interiores, que solo identifica la ubicación estática, el seguimiento de objetos proporciona una visión dinámica y detallada del recorrido de la persona, desde un punto A hasta un punto B, registrando su trayectoria y patrones de movimiento.

A lo largo de los años, diversos enfoques se han propuesto para afrontar los desafíos en el ámbito de la visión por computadora (Yilmaz *et al.*, 2006). En (Kim y Sim, 2010) se analiza el seguimiento de objetos específicos en entornos con múltiples objetos en movimiento, utilizando secuencias de video para identificar y rastrear el objetivo en cada cuadro. Este enfoque, parte de la generación de imágenes diferenciales mediante técnicas de seguimiento por regiones y emplea el filtro de partículas, logrando así un seguimiento robusto y preciso, incluso en escenarios complejos donde coexisten múltiples objetivos.

Por otra parte, en (Kim y Kweon, 2011) se aborda el desafío de la detección y el seguimiento de objetos mediante cámaras en movimiento, integrando técnicas basadas en homografía con el algoritmo de impulso de línea (*online-boosting*). Este método combina la extracción de características con algoritmos robustos, como KLT y RANSAC, para calcular la homografía y ajustar el seguimiento a condiciones dinámicas.

Otro enfoque relevante es el presentado en (Keivani *et al.*, 2017), donde se utiliza un algoritmo K-means modificado para la detección de múltiples objetos en movimiento en tiempo real. Este método opera de forma eficiente sin requerir información previa sobre el número de objetos en movimiento, adaptándose a escenarios dinámicos caracterizados por cambios en la iluminación o trayectorias no lineales.

A su vez, el trabajo de (Lychkov *et al.*, 2018) se destaca por las mejoras en la técnica del flujo óptico disperso, que permite un seguimiento más preciso en tiempo real mediante la regeneración automática de puntos característicos perdidos. Esta innovación resulta clave en situaciones donde los objetos enfrentan occlusiones, rotaciones o cambios de escala, incrementando la versatilidad del sistema.

En conjunto, estos trabajos reflejan la diversidad de enfoques en el ámbito de la visión por computadora, pero también evidencian limitaciones y desafíos por resolver, como la integración de estas técnicas en entornos cerrados o su aplicación específica al seguimiento de personas, particularmente adultos mayores.

En conclusión, el sistema propuesto en este trabajo se sitúa en la intersección de técnicas de seguimiento de objetos, métodos de localización en interiores y aplicaciones de monitoreo de adultos mayores. Si bien no ofrece una localización completa ni analiza el movimiento corporal detalladamente, permite obtener información relevante sobre la ubicación de una persona mayor dentro de un espacio controlado mediante una solución no intrusiva y fácilmente escalable.

3. Metodología

3.1. Modelo especializado

El modelo especializado en la detección y clasificación de adultos mayores se obtuvo mediante un proceso de cinco etapas, representado en la Figura 1. En esta sección, se presentan

de manera detallada el método de recolección de datos, las anotaciones, las estrategias de preprocesamiento, el proceso de entrenamiento del modelo y los métodos de validación aplicados para medir el desempeño del sistema.

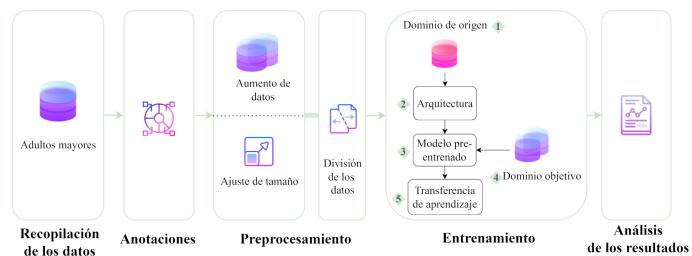


Figura 1: Flujo de trabajo para obtención y prueba de un modelo especializado.

3.1.1. Recopilación de los datos

La creación de un conjunto de datos especializado representó un factor determinante para el proyecto, lo que permitió el ajuste del modelo preentrenado (YOLOv8s) para la tarea de detección de adultos mayores. Durante esta etapa, se realizó una búsqueda de recursos en línea (por ejemplo: Pexels, Unsplash y Pixabay) con el fin de obtener imágenes adecuadas para el entrenamiento del modelo. Se seleccionaron imágenes que contenían a los adultos mayores en diferentes entornos, poses y condiciones de iluminación para asegurar diversidad y representatividad.

En esta fase, los desafíos principales fueron:

- La escasez de datos en una única fuente, lo que motivó la integración de imágenes de múltiples orígenes.
- El limitado material relacionado con adultos mayores.

En consecuencia, el conjunto de datos fue integrado por 2,418 imágenes de adultos mayores (60+ años) realizando actividades cotidianas como: caminar, correr, sentarse y estar de pie (ver Figura 2), donde predominaron las perspectivas parciales (tomas de un tercio o dos tercios del perfil).



Figura 2: Ejemplos de imágenes recopiladas de adultos mayores. Recuperado de (iStock, nd).

3.1.2. Anotaciones del conjunto de datos

El proceso de anotación se llevó a cabo utilizando cuadros delimitadores (*bounding boxes*) generados manualmente mediante la herramienta LabelMe (Russell et al., 2008; Sager et al., 2021), en formato COCO¹. Cada anotación incluyó: (1) las coordenadas de los vértices y (2) la etiqueta de la clase correspondiente (“0” para hombres y “1” para mujeres). Los datos se almacenaron en un formato JSON. Como resultado, se obtuvo una distribución de 1,210 etiquetas para cada clase (ver Figura 3).

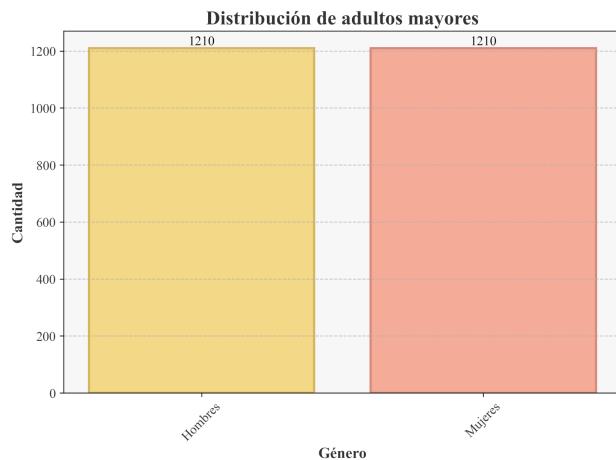


Figura 3: Distribución del conjunto de datos de adultos mayores.

3.1.3. Preprocesamiento

Durante la etapa de preprocesamiento, se aplicaron técnicas de aumento de datos con el fin de incrementar la variabilidad del conjunto antes del entrenamiento, mediante la incorporación de imágenes adicionales (negativas y con contexto) y transformaciones (geométricas y de color). Además, con el objetivo de asegurar que los datos cumplieran con los requisitos necesarios para su integración con el modelo preentrenado, se realizó un redimensionamiento de las imágenes. Finalmente, se hizo la división del conjunto de datos, lo que permitió entrenar, ajustar y evaluar el modelo de forma controlada, favoreciendo la generalización y evitando el sobreajuste.



Figura 4: Ejemplos de imágenes utilizadas en el conjunto de datos AM-2025.

Aumento de datos. Inicialmente, por cada imagen del conjunto de datos (Figura 4(a)) se extrajo una subimagen correspondiente a la región de interés (ROI) (ver Figura 4(b) y 4(c)) definida por los cuadros delimitadores (ver sección 3.1.2). Posterior-

mente, se aplicaron transformaciones geométricas y de color (ruido gaussiano, cambios de tono, ajustes de brillo, desenfoque, rotación, escalado, traslación y reflejos) en cada una de las subimágenes para obtener nueve variantes de cada una, ver Figura 5.



Figura 5: Transformaciones de color y geométricas.

Finalmente, se añadieron imágenes contextualizadas y negativas (*background*). Esta incorporación fue clave, ya que el modelo con el conjunto de prueba sólo detectaba la clase de interés pero no su ubicación dentro de la imagen. Como resultado, el conjunto de datos quedó integrado por 28,807 imágenes RGB distribuidas en tres categorías: (a) adultos mayores con contexto, (b)(c) adultos mayores sin contexto e (d) imágenes negativas.

Ajuste de tamaño. A continuación, se normalizaron las dimensiones de las imágenes. Para esto, se hizo uso de: (1) relleno (*padding*) para mantener la relación de aspecto original y (2) un redimensionamiento de 640×640 px para disminuir el consumo de recursos computacionales durante el entrenamiento. Al conjunto de datos se le denominó AM-2025 en referencia al grupo etario y al año de creación.

División del conjunto de datos. Por último, se empleó una partición tipo hold-out para separar los datos en conjuntos de entrenamiento (70 %), validación (15 %) y prueba (15 %), ver Tabla 1.

Tabla 1: Distribución de AM-2025.

Sin contexto	Con contexto	Brackground
24180	2209	2418

3.1.4. Entrenamiento

La etapa de entrenamiento mostrada en la Figura 6 amplía la visión general presentada en la Figura 1, al detallar las fases para la obtención de un modelo especializado aplicando transferencia de aprendizaje.

- ① Elección de un modelo preentrenado.
- ② Obtención de un conjunto de datos especializado.
- ③ Adaptación del modelo con TL.

¹Recuperado de <https://cocodataset.org/#format-data>

El TL se implementa cuando se dispone de una cantidad limitada de datos (Ramdan *et al.* (2020)), ya que permite reutilizar la información contenida en modelos preentrenados con datos genéricos más cuantiosos. Para ello, se entrena una red base y se reutilizan sus primeras capas en la arquitectura de la red de destino. Posteriormente, las capas restantes de la red de destino se inicializan aleatoriamente y se ajustan para adaptarse a la tarea específica (Yosinski *et al.* (2014)).

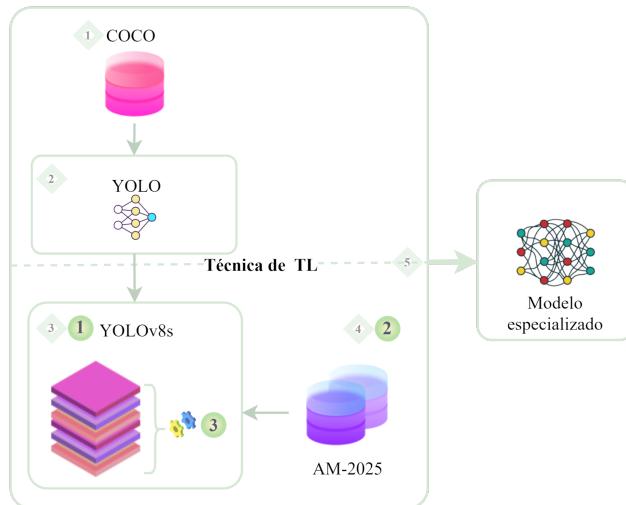


Figura 6: Fases en la transferencia de aprendizaje.

Para implementar la técnica de transferencia de aprendizaje utilizando el conjunto de datos AM-2025, se seleccionó la arquitectura de YOLO (Redmon *et al.*, 2016), utilizada en tareas de detección de objetos (Terven *et al.*, 2023). El modelo YOLOv8s (Jocher *et al.*, 2023) entrenado con el conjunto de datos COCO (Lin *et al.*, 2014) fue adaptado utilizando los hiperparámetros detallados en la Tabla 2.

Tabla 2: Hiperparámetros para el entrenamiento.

Parámetro	Descripción
Learning rate	Inicial (lr0): 0.01, Final (lrf): 0.1
Batch size	16 muestras por lote
Épocas	50 iteraciones completas
Optimizador	SGD con momentum (0.937)
Función de pérdida	$Loss_{total} = Loss_{cls} + Loss_{box} + Loss_{obj}$

Con los recursos computacionales descritos en la Tabla 3, el proceso de entrenamiento se completó en 9.35 h.

Tabla 3: Recursos computacionales.

Componente	Especificaciones
Hardware	NVIDIA GeForce RTX 4070
CPU	12th Gen Intel(R) Core(TM) i7 (2.3GHz)
Sistema operativo	Windows 11
	Python 3.10.16
Bibliotecas	Ultralytics 8.3.99 Pytorch 2.0.1+cu118 Sección 3.1.4 LabelMe 5.4.1
	OpenCV 4.11.0.8 Numpy 1.23.5 Sección 3.2

3.2. Estimación de trayectorias

El objetivo principal de este trabajo es garantizar la localización precisa y continua de adultos mayores en cualquier posición dentro de una imagen, con el fin de facilitar su seguimiento en espacios cerrados.

En este contexto, la predicción de trayectorias desempeña un papel crucial al reducir significativamente la probabilidad de perder el objetivo durante movimientos rápidos.

La implementación de un algoritmo propio (ver Figura 7) proporciona un control completo sobre su diseño y optimización, permitiendo adaptarlo de manera específica a las necesidades del proyecto. Además, facilita tanto el mantenimiento como la escalabilidad a medida que evolucionan los requisitos y se integra dentro de un sistema de mayor complejidad. Aunque algoritmos más elaborados pueden ofrecer un mejor rendimiento en escenarios complejos, su desarrollo puede demandar más tiempo y recursos computacionales.

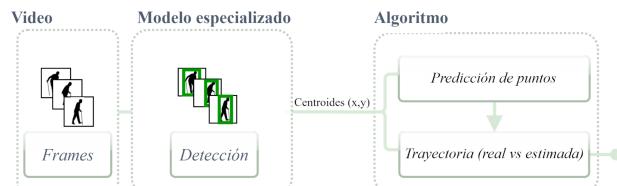


Figura 7: Flujo de trabajo para la estimación de trayectorias.

El Algoritmo 1 toma un par de puntos consecutivos y predice la ubicación de un nuevo punto siguiendo la misma dirección del desplazamiento entre puntos consecutivos. Esto permite extender la serie de puntos en la misma dirección en la que han estado cambiando. Mientras que el Algoritmo 2, toma la información generada por Algoritmo 1 y construye una trayectoria real y una trayectoria estimada.

Algoritmo 1 Predicción de puntos

Entrada: Dos puntos $A(x_1, y_1)$ y $B(x_2, y_2)$ que definen el segmento.

Salida: Un tercer punto $C(x_3, y_3)$ alineado con la dirección de la recta y a la misma distancia que AB .

1: **Vector de dirección:**

2: $\Delta_x = x_2 - x_1$

3: $\Delta_y = y_2 - y_1$

4: **Obtención del punto C:**

5: $x_3 = x_2 + \Delta_x$

6: $y_3 = y_2 + \Delta_y$

7: **Salida:** $C(x_3, y_3)$ **Algoritmo 2** Trayectoria

Entrada: Lista de centroides reales **centroides_r**.

Salida: Lista de puntos proyectados **puntos_p**.

Iniciar puntos_p como una lista vacía.

2: **para** cada centroide en **centroides_r** **hacer**

si no es el último centroide **entonces**

 4: Obtener a como el centroide actual.

 Obtener b como el siguiente centroide.

 6: $c_{\text{proyectado}} \leftarrow$ Basado en a y b (Algoritmo 1).

 Agregar $c_{\text{proyectado}}$ a **puntos_p**.

8: **fin si**

fin para

4. Resultados

Esta sección presenta los resultados en dos fases: (1) el entrenamiento, validación y evaluación del modelo especializado, y (2) la aplicación y evaluación del algoritmo de estimación de trayectorias.

4.1. Modelo especializado**4.1.1. Entrenamiento**

En las gráficas presentadas en la Figura 8, se observa la evolución del modelo durante el entrenamiento. Los resultados confirman que tanto la pérdida asociada a los cuadros delimitadores (*bounding box loss*, *box_loss*), como la pérdida de clasificación (*classification loss*, *cls_loss*) convergen rápidamente hacia valores próximos a cero, lo que indica estabilidad en el proceso de aprendizaje.

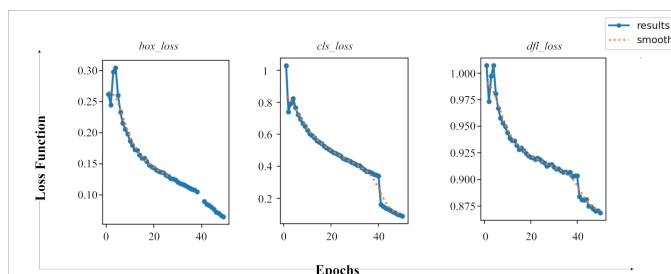


Figura 8: Pérdidas de entrenamiento (*training loss*).

4.1.2. Validación

En la Figura 9, se confirma que el modelo ha aprendido a localizar e identificar correctamente a los adultos mayores en el conjunto de validación. Se tiene precisión en la ubicación y tamaño de las personas al tener valores de pérdida bajas y estables (*box_loss*). Además, se mantiene una discriminación efectiva entre clases al presentar una tasa de error que converge rápidamente hacia cero (*cls_loss*).

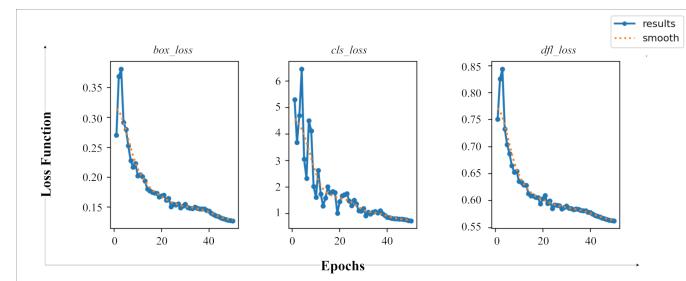


Figura 9: Pérdidas de validación (*validation loss*).

En la Figura 10, se observa la evolución del desempeño del modelo a lo largo del tiempo frente a datos desconocidos.

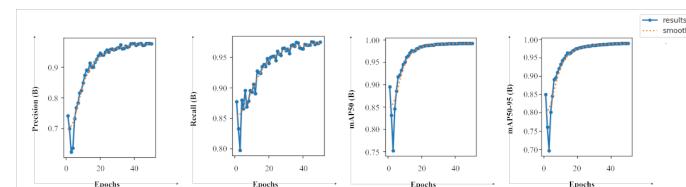


Figura 10: Métricas de desempeño del modelo con los datos de validación.

En la Tabla 4 se resume el rendimiento del modelo, con métricas de precisión (*precision*, P), sensibilidad (*recall*, R) y precisión media promedio (*mean average precision*, mAP) cercanas a 1, lo que sugiere una alta confiabilidad en la detección y clasificación de las personas. Las pérdidas ABL y ACL confirman un entrenamiento robusto, con errores mínimos en localización y clasificación.

Tabla 4: Métricas de validación.

P	R	mAP50	mAP50-95	ABL	ACL
0.976	0.975	0.992	0.989	0.147	0.448

Nota: ABL = pérdida promedio de las cajas, ACL= pérdida promedio por clase.

A pesar de que las imágenes negativas no contribuyen a las métricas de las clases de interés, los falsos positivos en el *background* se reflejan en la matriz de confusión como predicciones incorrectas del modelo, ver Figura 11.

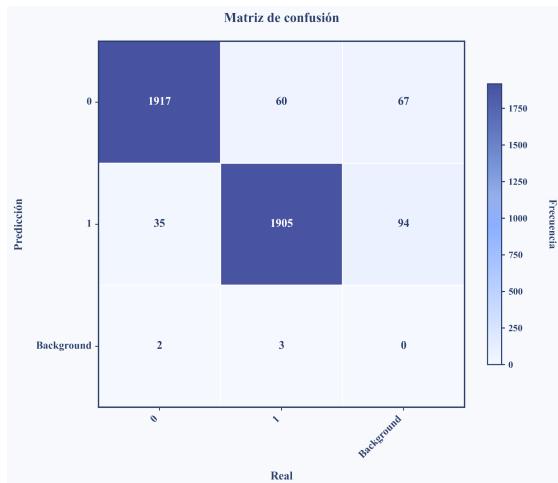


Figura 11: Matriz de confusión.

Las métricas presentadas en la Tabla 5 corresponden a los resultados de las dos clases de interés.

Tabla 5: Métricas por clase.

Clase	P	R	F1-Score
0	0.982	0.969	0.975
1	0.969	0.982	0.975

4.1.3. Prueba

Utilizando el conjunto de prueba se evaluó el rendimiento final del modelo. En la Tabla 6 se presentan los resultados obtenidos y se observa la prevalencia en el desempeño.

Tabla 6: Métricas generales.

P	R	mAP50	mAP50-95
0.9841	0.9613	0.9788	0.9758

Así mismo, la matriz de confusión (ver Figura 12) y las métricas presentadas en la Tabla 7 muestran que el modelo alcanza un rendimiento robusto y equilibrado, con errores menores al 0.02 en ambas clases.

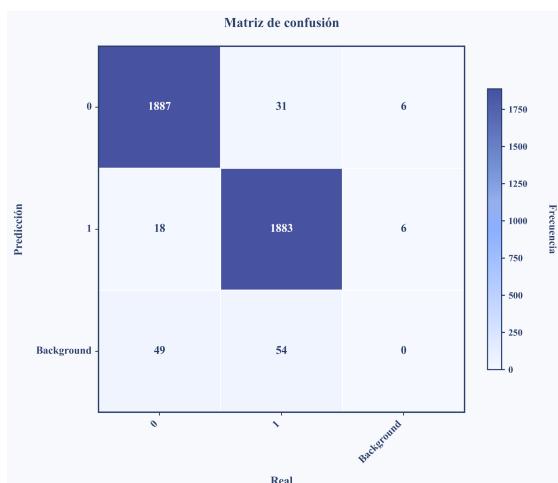


Figura 12: Matriz de confusión.

Tabla 7: Métricas por clase.

Clase	P	R	F1-Score
0	0.990	0.983	0.987
1	0.983	0.980	0.987

En la Figura 13, se muestran ejemplos de las predicciones realizadas sobre imágenes externas al conjunto de entrenamiento y validación. Las imágenes contienen: 1) la ubicación de los adultos mayores por medio de cuadros delimitadores, 2) las etiquetas para cada una de las clases y 3) la precisión del modelo asociada a cada predicción.

● Mujer [1] ● Hombre [0]

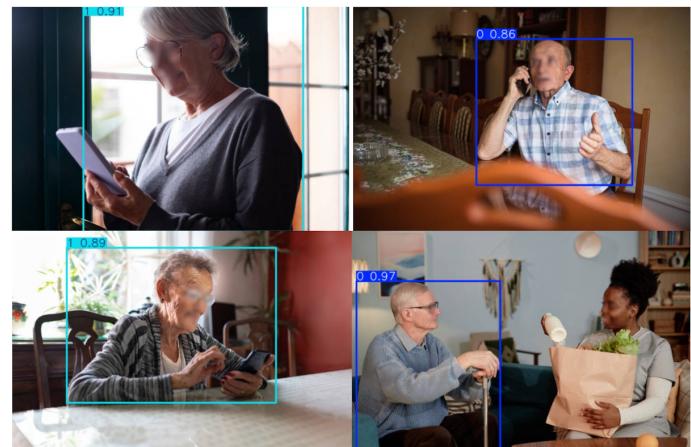


Figura 13: Ejemplos de predicciones generadas utilizando el modelo especializado en el conjunto de prueba.

4.2. Algoritmo de seguimiento

Con el objetivo de analizar el rendimiento del modelo en tareas de seguimiento se llevaron a cabo evaluaciones utilizando videos cortos (≈ 15 s). Los resultados alcanzados por el algoritmo de estimación de trayectorias cumplen con las expectativas establecidas. En las Figuras 14 y 15 se muestran secuencias de frames junto con una vista detallada de la trayectoria real de los adultos mayores.

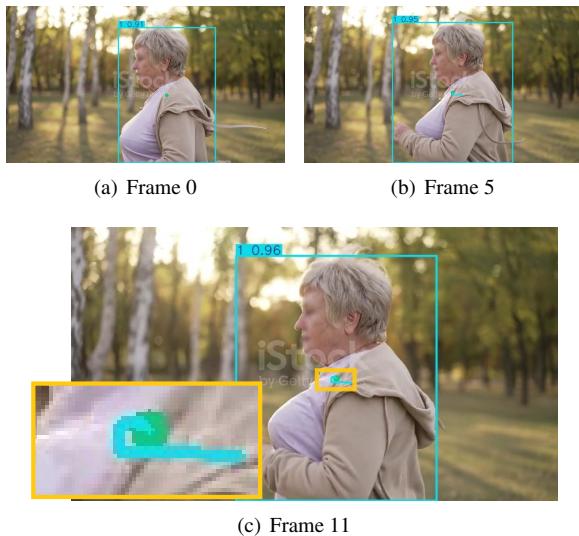


Figura 14: Prueba 1 para estimación de la trayectoria.

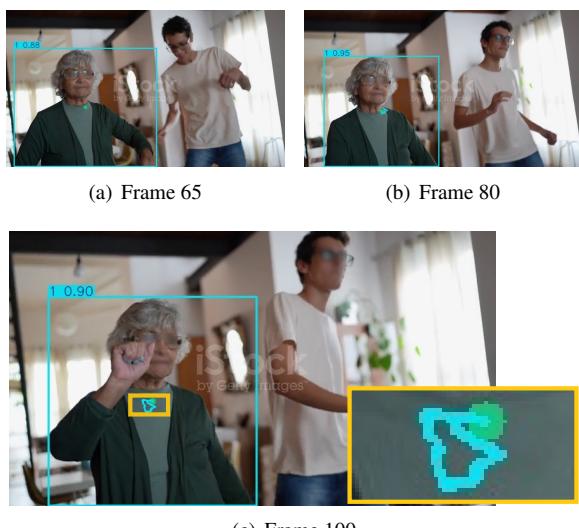


Figura 15: Prueba 2 para estimación de la trayectoria.

En las Figuras 16 y 17, se presenta una comparativa entre la trayectoria real y la estimada a través de los cálculos realizados por el algoritmo propuesto.

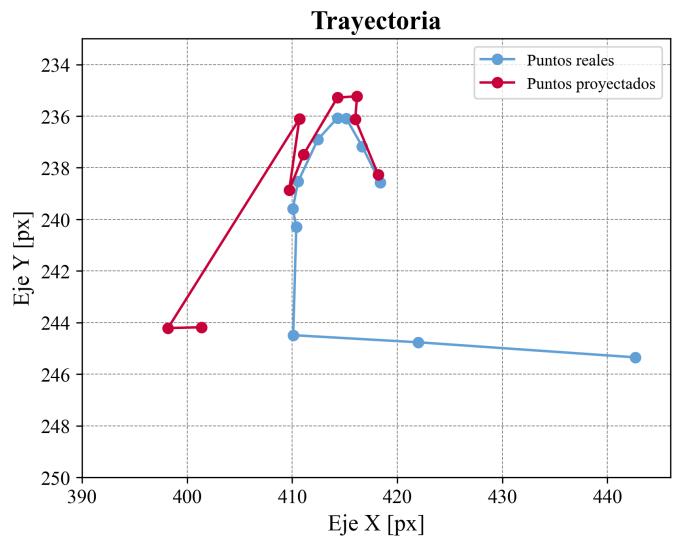


Figura 16: Estimación de la trayectoria (Prueba 1).

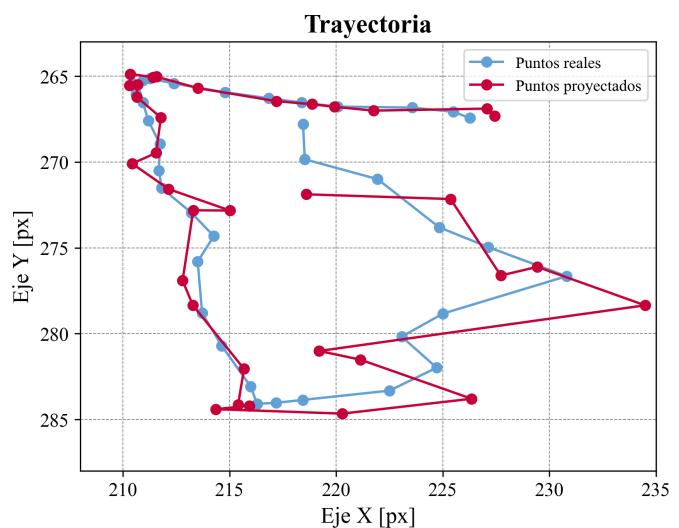


Figura 17: Estimación de la trayectoria (Prueba 2).

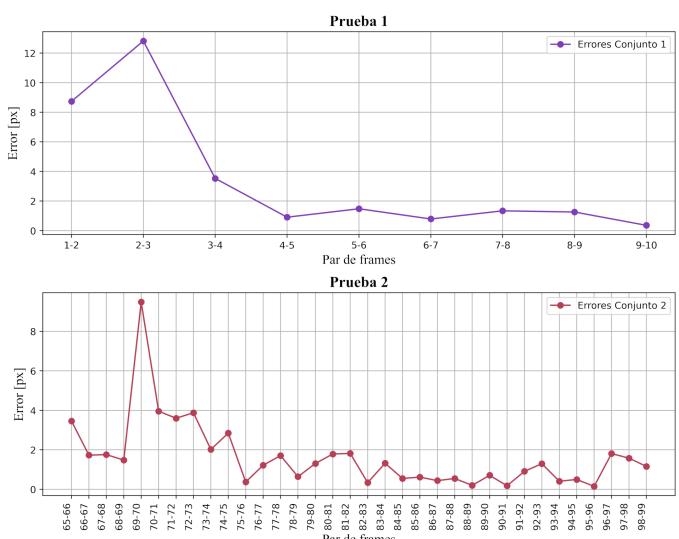


Figura 18: Errores de distancia entre pares de frames consecutivos.

Los valores de error que se muestran en la Figura 18 denotan una disminución significativa conforme avanza la secuencia de mediciones. Inicialmente, los errores son relativamente altos (en la Prueba 1 inicia con 8.74 px y 3.46 px en la Prueba 2) debido a condiciones iniciales menos controladas. Sin embargo, a partir del par de frames 3-4 se observa una tendencia decreciente y una estabilización progresiva de los valores, alcanzando cifras mínimas; por ejemplo 0.36 px en el frame 9-10 y 1.16 px en el frame 98-99. Esta evolución sugiere que el algoritmo logra mejorar su desempeño conforme se ajusta a los datos, mostrando así una mayor precisión en las etapas posteriores.

En la sección 4.1, el modelo especializado alcanzó un desempeño general de 0.989 en mAP50-95. Esto se vio reflejado en las Figuras 13, 14 y 15, donde se observó su comportamiento frente a dos condiciones: (a) adultos mayores sin compañía y (b) en compañía de jóvenes, así como en dos entornos: externos e internos. Este resultado sirvió como base para la tarea de seguimiento (sección 4.2), en el que se analizó el margen de error (con un máximo de 13 px en la Prueba 1 y 9 px en la Prueba 2) entre la trayectoria real y la estimada por el algoritmo propuesto para obtener el historial del recorrido del adulto mayor dentro del campo de visión.

5. Discusión y conclusiones

En este trabajo, se utilizó un conjunto de datos enfocado en personas de la tercera edad (AM2025) para aplicar la técnica de transferencia de aprendizaje a un modelo preentrenado (YOLOv8s). Como resultado, el modelo especializado obtuvo un mAP50-95 de 0.989, considerando la precisión en la ubicación y clasificación en distintos niveles. Además, los resultados de las Figuras 16 y 17 demuestran que el algoritmo implementado para la estimación de trayectorias es capaz de obtener un histórico del recorrido del adulto mayor dentro del campo de visión.

En términos prácticos, la aplicación de un modelo especializado podría permitir la implementación de un sistema no intrusivo para el seguimiento de personas de la tercera edad en entornos cerrados (hogares o asilos), facilitando un monitoreo personalizado.

Estos hallazgos están en línea con investigaciones previas como las descritas en (Karkar *et al.*, 2021), (Chen, 2022), (Chen *et al.*, 2022), (Lee *et al.*, 2024) y (Mane *et al.*, 2025) donde los sistemas desarrollados se enfocan en personas de la tercera edad. Sin embargo, las tecnologías utilizadas no permiten distinguir con precisión si se trata de una persona de la tercera edad ni tampoco el género.

El algoritmo de estimación de trayectorias basado en un modelo especializado en adultos mayores demostró un funcionamiento adecuado en los casos analizados. La precisión alcanzada permite considerar su aplicación en sistemas de mayor complejidad, tales como plataformas robóticas con módulos de visión artificial para tareas de asistencia personalizada.

Agradecimientos

Este trabajo fue posible gracias al apoyo de la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI), a través de la Convocatoria de Becas Nacionales para Estudios de Posgrado.

Referencias

- Ahmed, H. M. y Abdulrazak, B. (2021). Monitoring indoor activity of daily living using thermal imaging: A case study. *International Journal of Advanced Computer Science and Applications*, 12(9).
- Chen, M.-Y. (2022). Establishing a cybersecurity home monitoring system for the elderly. *IEEE Transactions on Industrial Informatics*, 18(7):4838–4845.
- Chen, Y., Zhang, Y., Xiao, B., y Li, H. (2022). A framework for the elderly first aid system by integrating vision-based fall detection and bim-based indoor rescue routing. *Advanced Engineering Informatics*, 54:101766.
- Chifu, V. R., Pop, C. B., Demjen, D., Socaci, R., Todea, D., Antal, M., Cioara, T., Anghel, I., y Antal, C. (2022). Identifying and monitoring the daily routine of seniors living at home. *Sensors*, 22(3).
- Deepthi, P. (2025). Deep learning in video surveillance: Enhancing security and threat detection. *International Journal of Scientific Research & Engineering Trends*, 11(2):2086–2090. Volume 11, Issue 2.
- Fuentes, D., Correia, L., Costa, N., Reis, A., Ribeiro, J., Rabadão, C., Barroso, J., y Pereira, A. (2021). Indoocare: Low-cost elderly activity monitoring system through image processing. *Sensors*, 21(18).
- Georgoulas, C., Linner, T., y Bock, T. (2014). Towards a vision controlled robotic home environment. *Automation in Construction*, 39:106–116.
- Güttler, J., Georgoulas, C., Linner, T., y Bock, T. (2014). Towards a Future Robotic Home Environment: A Survey. *Gerontology*, 61(3):268–280.
- iStock (n.d.). istock - fotos, imágenes y videos de stock. <https://www.istockphoto.com/>. Recuperado en octubre 2024.
- Jocher, G., Chaurasia, A., y Qiu, J. (2023). Ultralytics yolov8.
- Karkar, A. G., Al-Maadeed, S., Kunthot, J., y Bouridane, A. (2021). Camnav: a computer-vision indoor navigation system. *The Journal of Supercomputing*, 77(7):7737–7756.
- Keivani, A., Tapamo, J.-R., y Ghayoor, F. (2017). Motion-based moving object detection and tracking using automatic k-means. En *2017 IEEE AFRICON*, pp. 32–37.
- Kim, D., Bian, H., Chang, C., Dong, L., y Margrett, J. (2022). In-home monitoring technology for aging in place: Scoping review. *Interactive Journal of Medical Research*, 11(2):e39005.
- Kim, H.-B. y Sim, K.-B. (2010). A particular object tracking in an environment of multiple moving objects. En , pp. 1053–1056.
- Kim, W. J. y Kweon, I.-S. (2011). Moving object detection and tracking from moving camera. En *2011 8th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, pp. 758–759.
- Lee, C., Chu, E. T.-H., Sie, M., Lin, L., Hong, M., y Huang, C. (2024). Application of indoor positioning systems in nursing homes: Enhancing resident safety and staff efficiency. *Sensors*, 24(18):6099.
- Lin, T., Maire, M., Belongie, S. J., Bourdev, L. D., Girshick, R. B., Hays, J., Perona, P., Ramanan, D., Dollár, P., y Zitnick, C. L. (2014). Microsoft COCO: common objects in context. *CoRR*, abs/1405.0312.
- Lychkov, I. I., Alfimtsev, A. N., y Sakulin, S. A. (2018). Tracking of moving objects with regeneration of object feature points. En *2018 Global Smart Industry Conference (GloSIC)*, pp. 1–6.
- Mane, V., Mahajan, R., Durge, H. A., Ghosh, P., Kadam, K., Mahajan, I., y Muneshwar, P. (2025). Alert system for non-responsive state of an elderly person. *Journal of Computational Systems and Applications*. Corresponding author: Vijay Mane, Vijay.mane@vit.edu.
- Nayak, R., Pati, U. C., y Das, S. K. (2020). A comprehensive review on deep learning-based methods for video anomaly detection. *Image and Vision Computing*, 104078:104078. Received: 28 August 2020; Revised: 12 November 2020; Accepted: 25 November 2020.
- Nguyen, L. N., Susarla, P., Mukherjee, A., Cañellas, M. L., Álvarez Casado, C., Wu, X., Silvén, O., Jayagopi, D. B., y López, M. B. (2024). Non-contact multimodal indoor human monitoring systems: A survey. *Information Fusion*, 110:102457.
- Organización Mundial de la Salud (2015). Informe mundial sobre el envejecimiento y la salud. Informe técnico. Disponible en <https://iris.who.int/handle/10665/186466>.
- Organización Mundial de la Salud (2024). Ageing and health. Último acceso: 27 de marzo de 2025.
- Ramdan, A., Heryana, A., Arisal, A., Kusumo, R. B. S., y Pardede, H. F. (2020). Transfer learning and fine-tuning for deep learning-based tea diseases detection on small datasets. En *2020 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET)*, pp. 206–211.
- Redmon, J., Divvala, S., Girshick, R., y Farhadi, A. (2016). You only look once: Unified, real-time object detection. En *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788.

- Russell, B. C., Torralba, A., Murphy, K. P., y Freeman, W. T. (2008). Labelme: A database and web-based tool for image annotation. *Int. J. Comput. Vision*, 77(1–3):157–173.
- Sager, C., Janiesch, C., y and, P. Z. (2021). A survey of image labelling for computer vision applications. *Journal of Business Analytics*, 4(2):91–110.
- Sawik, B., Tobis, S., Baum, E., Suwalska, A., Kropińska, S., Stachnik, K., Pérez-Bernabeu, E., Cildoz, M., Agustín, A., y Wieczorowska-Tobis, K. (2023). Robots for elderly care: Review, multi-criteria optimization model and qualitative case study. *Healthcare (Basel)*, 11(9):1286.
- Terven, J., Córdova-Esparza, D.-M., y Romero-González, J.-A. (2023). A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas. *Machine Learning and Knowledge Extraction*, 5(4):1680–1716.
- Vélez, Z., Fabián, T., Jamileth, T., Álvarez, L., Saverio, T., Jamileth, C., Alencastro, P., y Alexander, J. (2024). Avances en la tecnología asistiva para mejorar la calidad de vida en adultos mayores. revisión actualizada en el contexto latinoamericano. *Ciencia Latina Revista Científica Multidisciplinar*, 8(1):645–662.
- Yilmaz, A., Javed, O., y Shah, M. (2006). Object tracking: a survey. *ACM Computing Surveys (CSUR)*, 38(4):1–45. Source: DBLP.
- Yosinski, J., Clune, J., Bengio, Y., y Lipson, H. (2014). How transferable are features in deep neural networks? *Advances in Neural Information Processing Systems (NIPS)*, 27.
- Zin, T. T., Htet, Y., Akagi, Y., Tamura, H., Kondo, K., Araki, S., y Chosa, E. (2021). Real-time action recognition system for elderly people using stereo depth camera. *Sensors*, 21(17).