

Embedded Computer Vision Safety System for Freight Elevators Using SSD-MobileNet

Yojar Dudayev Apaza Chullunquia, Degree¹, Cristhian Jair Guzman Huaman, Degree², Alert Mendoza Acosta, Mag.³, Edward Russel Sanchez Penadillo, Mag.⁴
^{1,2,3,4}Universidad Tecnológica del Perú, Perú, u17300711@utp.edu.pe, u19305070@utp.edu.pe, e08141@utp.edu.pe, e10173@utp.edu.pe

Abstract-- This paper presents the design and implementation of an intelligent safety system for freight elevators based on computer vision, aimed at reducing workplace accidents caused by improper use of such equipment. The proposed system relies on an SSD-MobileNet convolutional neural network, trained with a dataset of 1,050 labeled images under varying lighting conditions and deployed on a low-cost ESP32-CAM microcontroller. The system detects the presence of individuals at the elevator entrance and, through communication with a Siemens PLC S7-1200 and a variable frequency drive (VFD), determines whether to enable or block motor activation. Validation was conducted in a controlled laboratory environment using a three-level platform with a 25 kg load. The experimental results yielded an F1-score of 93.13%, a recall of 90.63%, and a specificity of 96.00%. The system is presented as a functional proof of concept with future potential for deployment in real industrial environments, highlighting its low cost, effective integration, and preventive approach to occupational safety.

Keywords-- Computer vision, Freight elevator safety, SSD-MobileNet, Embedded systems, Occupational safety

Sistema de Seguridad con Visión Artificial Embebida para Ascensores de Carga Usando SSD-MobileNet

Yojar Dudayev Apaza Chullunquia, Degree¹, Cristhian Jair Guzman Huaman, Degree², Alert Mendoza Acosta, Mag.³, Edward Russel Sanchez Penadillo, Mag.⁴
^{1,2,3,4}Universidad Tecnológica del Perú, Perú, u17300711@utp.edu.pe, u19305070@utp.edu.pe, e08141@utp.edu.pe, e10173@utp.edu.pe

Resumen– Este artículo presenta el diseño e implementación de un sistema inteligente de seguridad para ascensores de carga basado en visión artificial, orientado a reducir accidentes laborales ocasionados por el uso indebido de dichos equipos. El sistema propuesto utiliza una red neuronal convolucional SSD-MobileNet, entrenada con un conjunto de 1,050 imágenes etiquetadas bajo condiciones variables de iluminación, e implementada en un microcontrolador de bajo costo ESP32- CAM. El sistema detecta la presencia de personas en el acceso al ascensor y mediante comunicación con un PLC Siemens S7- 1200 y un variador de frecuencia (VFD), determina si se habilita o bloquea la activación del motor. La validación se realizó en un entorno controlado de laboratorio utilizando una plataforma de tres niveles con una carga de 25 kg. Los resultados experimentales arrojaron un F1-score de 93.13 %, un recall de 90.63 % y una especificidad de 96.00 %. El sistema se presenta como una prueba de concepto funcional, con potencial de implementación futura en entornos industriales reales, destacando por su bajo costo, integración efectiva y enfoque preventivo en seguridad ocupacional.

Palabras clave-- *Visión artificial, Seguridad en ascensores de carga, SSD-MobileNet, Sistemas embebidos, Seguridad ocupacional.*

I. INTRODUCCIÓN

En entornos industriales, los elevadores de carga son esenciales para el transporte vertical de materiales. Sin embargo, el uso indebido por parte de personas representa un riesgo significativo de accidentes laborales, especialmente cuando se activan sin respetar los protocolos de seguridad. Según el Ministerio de Trabajo y Promoción del Empleo (MTPE), se han registrado múltiples accidentes asociados al uso inadecuado de plataformas elevadoras, destacando la necesidad de sistemas de seguridad más eficientes [1].

Las soluciones tradicionales para prevenir estos incidentes, como sensores mecánicos, fotoceldas o pulsadores de emergencia, presentan limitaciones en entornos reales, donde factores como la iluminación, el polvo o la manipulación deliberada pueden comprometer su funcionamiento [2]. En los últimos años, la visión artificial ha emergido como una alternativa eficaz para la detección de personas en espacios controlados, gracias a los avances en

redes neuronales convolucionales (CNNs), las cuales permiten identificar objetos con alta precisión en tiempo real [3], [4].

Entre las arquitecturas más eficientes para sistemas embebidos destaca SSD-MobileNet, una red optimizada para dispositivos con recursos limitados, capaz de mantener un balance entre precisión y velocidad de inferencia [5], [6]. Su implementación en módulos como el ESP32-CAM, que integra cámara y procesamiento en un solo dispositivo de bajo costo, ha permitido desarrollar soluciones de monitoreo visual en tiempo real aplicadas en robótica, vigilancia y automatización industrial [7], [8].

A pesar del potencial de estas tecnologías, la mayoría de las investigaciones recientes se han limitado a tareas de clasificación o monitoreo sin intervención activa sobre el sistema de control [9], [10]. La integración directa con sistemas de automatización industrial, como los controladores lógicos programables (PLCs) y variadores de frecuencia, sigue siendo un desafío poco abordado [11], [12]. Además, se requiere validar este tipo de soluciones en escenarios cercanos a las condiciones reales de operación: variabilidad de iluminación, ruido electromagnético, interferencias físicas y tiempos de respuesta críticos [13], [14].

Este trabajo tiene como objetivo diseñar un sistema inteligente de seguridad para elevadores de carga, basado en visión artificial embebida mediante un ESP32-CAM ejecutando una red SSD-MobileNet para la detección de personas en el acceso al elevador. Esta detección se comunica en tiempo real con un PLC Siemens S7-1200, encargado de activar o bloquear un variador de frecuencia (VFD) que controla el motor del elevador. El desarrollo se estructura bajo el modelo en cascada y se valida en un prototipo funcional de tres niveles con carga real. A diferencia de trabajos anteriores, este sistema implementa visión artificial en tiempo real, transmisión serial segura, interbloqueo físico con PLC, y pruebas experimentales con métricas cuantitativas que demuestran su viabilidad.

II. METODOLOGIA

La metodología adoptada para el desarrollo del sistema de seguridad se estructuró bajo el enfoque de diseño en cascada, el cual permitió una implementación progresiva y verificable, considerando desde el análisis del problema hasta las pruebas funcionales. Las etapas consideradas fueron: análisis, diseño conceptual, desarrollo del modelo de visión artificial, integración con el sistema de control, implementación del sistema embebido y validación experimental.

A. Análisis de problema

El uso indebido de elevadores de carga por parte de personas representa una problemática frecuente en instalaciones industriales. La necesidad de un sistema de supervisión automatizado que identifique la presencia humana en zonas no autorizadas motivó el desarrollo de un sistema basado en visión artificial que actúe antes de iniciar el ciclo de operación del elevador.

Se definieron los siguientes requisitos funcionales:

- Detectar personas en el acceso del elevador en tiempo real.
- Activar un sistema de interbloqueo que evite el arranque si hay presencia humana.
- Registrar el evento para diagnóstico posterior.

B. Diseño del sistema (arquitectura general)

El sistema se compone de tres niveles funcionales:

- Nivel de percepción (visión embebida): encargado de capturar e interpretar imágenes del entorno.
- Nivel lógico (control embebido): encargado de procesar la información de visión y emitir señales de control.
- Nivel de actuación (PLC + VFD): responsable de ejecutar o bloquear el movimiento del elevador.

La arquitectura del sistema propuesto se resume en la Fig. 1, que muestra el flujo de señales desde el módulo de visión hasta el actuador del elevador, pasando por la lógica embebida y el PLC.

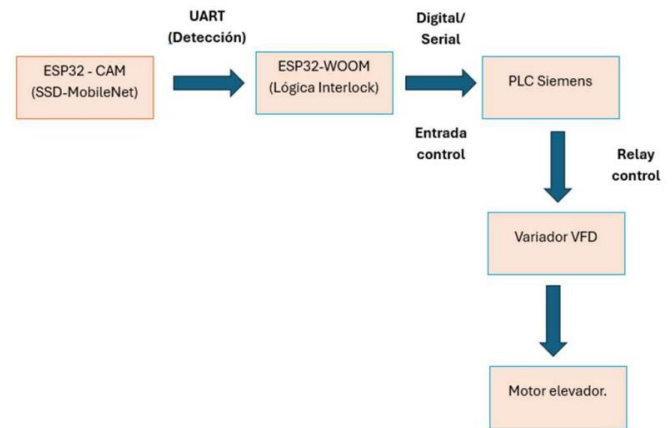


Fig. 1 Diagrama de bloques funcional del sistema propuesto.

C. Modelo de detección con SSD-MobileNet

Se eligió la arquitectura SSD-MobileNetV2 por su bajo costo computacional, buena precisión en detección de objetos pequeños, y capacidad de ejecución en tiempo real en dispositivos embebidos como el ESP32-CAM. A diferencia de arquitecturas más pesadas como YOLOv5 o Faster R-CNN, esta red mantiene un balance ideal entre eficiencia y desempeño.

- Dataset: se creó un conjunto de 1,050 imágenes etiquetadas en dos clases: “persona presente” y “sin persona”, capturadas en el entorno de prueba, variando iluminación, fondo y perspectiva.
- Preprocesamiento: resizing a 224x224 px, normalización, y data augmentation (rotación, brillo, zoom, espejo).
- Entrenamiento: se usó TensorFlow en Google Colab, con 50 épocas, tasa de aprendizaje 0.001, y batch size 32. El modelo final fue exportado como TFLite.
- Métricas de evaluación: accuracy, precision, recall, F1-score, y curva ROC.

D. Implementación embebida en ESP32-CAM

El modelo fue cargado en el ESP32-CAM utilizando la librería ESP-DL. La inferencia se ejecuta cada 1.5 segundos. Cuando se detecta una persona, el módulo genera un mensaje de alerta que se transmite mediante protocolo UART a un segundo microcontrolador ESP32-WROOM, encargado de aislar la lógica de control del procesamiento visual.

E. Integración con el PLC y variador

El ESP32-WROOM se comunica con el PLC Siemens PLC S7- 1200 mediante señal digital (emulando una entrada discreta) o UART, según la configuración de prueba. El PLC

analiza esta entrada junto con las condiciones del sistema, y decide si permite el arranque del motor trifásico mediante el variador de frecuencia Delta VFD075E23A. Si el PLC recibe una señal de “presencia humana”, interrumpe el ciclo de arranque por seguridad.

F. Verificación funcional y validación experimental

La validación se desarrolló sobre un prototipo de tres niveles con capacidad de carga de 25 kg. Se ejecutaron más de 50 ciclos de prueba, bajo las siguientes condiciones:

- Iluminación natural y artificial, con y sin contraluz
- Simulación de presencia con personas de distintas estaturas
- Pruebas de latencia de detección y bloqueo
- Cálculo de métricas: precisión, sensibilidad, especificidad, F1-score y tiempo de respuesta

Se registró un tiempo de respuesta promedio de 236 ms, dentro del umbral requerido para bloquear la activación del variador antes del movimiento del motor. La precisión global del modelo fue del 95.78%, con un F1-score del 93.13%, demostrando que el sistema puede operar de forma fiable en condiciones simuladas.

Finalmente, la integración de los módulos descritos permite el funcionamiento sincrónico del sistema. La cámara ESP32-CAM detecta la presencia de personas, el ESP32-WROOM gestiona los relés, y el PLC Siemens coordina las decisiones de control. La HMI permite visualizar el estado del ascensor.

La Fig. 2 presenta el flujo funcional completo del sistema embebido, desde la captura de imágenes hasta la ejecución del bloqueo del motor.

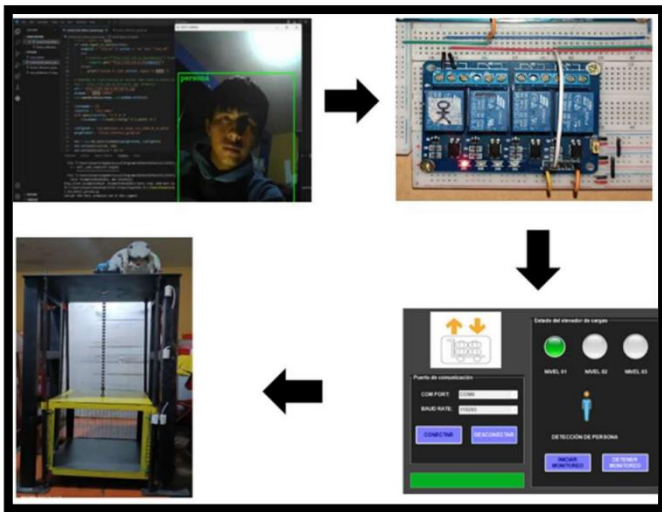


Fig. 2 Flujo funcional del sistema embebido: desde la detección hasta el control del motor.

III. RESULTADOS

El sistema propuesto fue validado experimentalmente en un entorno simulado controlado, replicando las condiciones típicas de operación de un elevador de carga de tres niveles. La evaluación se centró en dos aspectos principales: (a) el desempeño del modelo de detección basado en visión artificial y (b) el funcionamiento integral del sistema embebido junto con el controlador lógico programable (PLC) y el variador de frecuencia (VFD).

A. Desempeño del modelo de detección (SSD-MobileNet)

El módulo de visión artificial fue evaluado mediante un conjunto de pruebas con 1600 imágenes (800 con presencia de persona y 800 sin ella), capturadas bajo diferentes ángulos e intensidades de iluminación. El modelo SSD-MobileNet implementado en el microcontrolador ESP32-CAM demostró una capacidad efectiva de detección en tiempo real, con una tasa de inferencia aproximada de 5 cuadros por segundo (fps).

Las métricas obtenidas durante la validación del modelo fueron las siguientes: exactitud (accuracy) de 93.31 %, precisión positiva de 95.78 %, sensibilidad (recall) de 90.63 %, especificidad de 96.00 %, F1-Score de 93.13 % y una tasa de falsos positivos de apenas 4 %. Estos resultados evidencian un buen equilibrio entre precisión y sensibilidad, lo cual es fundamental en sistemas de seguridad, donde tanto los falsos negativos como los falsos positivos pueden comprometer la operatividad.

La Figura 3 muestra la matriz de confusión obtenida a partir del modelo entrenado, donde se aprecia que el sistema clasificó correctamente 725 imágenes con personas y 768 imágenes sin personas, con apenas 32 falsos positivos y 75 falsos negativos.

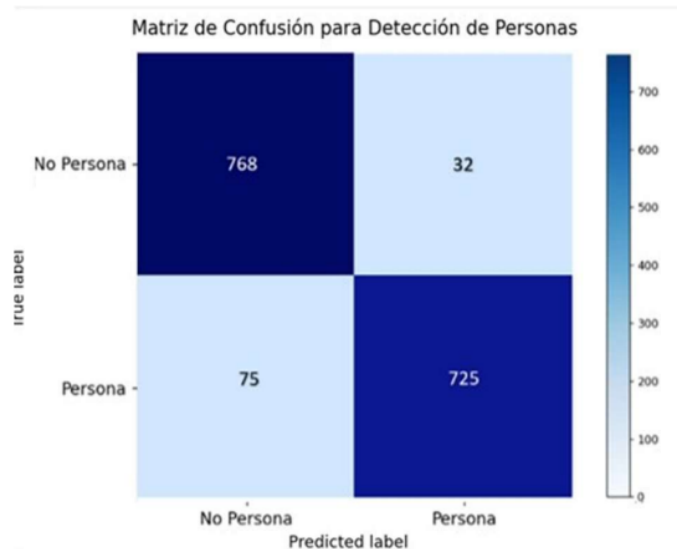


Fig. 2 Matriz de confusión del modelo SSD-MobileNet en la etapa de Validación.

Además, la siguiente tabla resume los principales indicadores del modelo, permitiendo una visualización comparativa del desempeño general alcanzado durante la etapa de validación.

TABLA I
MÉTRICAS DE DESEMPEÑO

Métrica	Valor (%)
Precisión	95.99
Exactitud	95.78
Sensibilidad	90.63
Especificidad	95.07
F1-Score	93.13

Estas métricas evidencian un desempeño robusto del modelo para la clasificación binaria de presencia humana, manteniendo un equilibrio adecuado entre tasa de verdaderos positivos y tasa de falsos positivos.

B. Desempeño del modelo de detección (SSD-MobileNet)

La respuesta del sistema ante la detección de una persona en la zona de acceso del ascensor fue evaluada mediante la medición de los tiempos involucrados en el ciclo completo: detección, procesamiento y activación del relé de seguridad. El tiempo total de respuesta es de aproximadamente 240 milisegundos, compuesto por:

- Inferencia del modelo SSD-MobileNet: 200 ms
- Lógica antirrebote (debounce): 30 ms
- Tiempo de escaneo del PLC: 10 ms

Este valor de latencia es determinante para garantizar que el sistema pueda detener el ascensor antes de que se produzca un desplazamiento peligroso. En condiciones de velocidad nominal del elevador (0.4 m/s), el desplazamiento de este durante este tiempo sería de tan solo 0.096 metros (9.6 cm), una distancia considerada segura dentro del margen de operación del sistema.

La Tabla 2 muestra el desplazamiento de la cabina en función de distintos valores de latencia y velocidades típicas de operación. Esta información es crítica para establecer los umbrales de seguridad y verificar que el sistema cumple con los tiempos máximos permitidos para la detención segura del ascensor.

TABLA II
DISTANCIA DE DESPLAZAMIENTO DE LA CABINA SEGÚN LA LATENCIA DEL SISTEMA.

Latencia (ms)	Vel. 0.4 m/s (m)	Vel. 0.6 m/s (m)	Vel. 0.8 m/s (m)
100	0.04	0.06	0.08
200	0.08	0.12	0.16
300	0.12	0.18	0.24

Para facilitar la interpretación visual de estos valores, la Fig. 4 presenta la relación entre la latencia total del sistema y la distancia de desplazamiento de la cabina del ascensor, según distintas velocidades de operación. Esta gráfica refuerza la conclusión de que la respuesta del sistema cumple con los requerimientos de seguridad incluso en escenarios de velocidad moderadamente elevada.

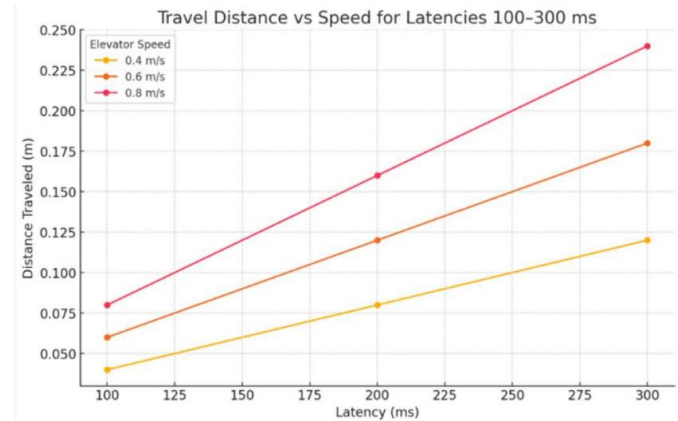


Fig. 4. Distancia recorrida vs velocidad del elevador

C. Actuación del sistema de control (PLC, relés y variador de frecuencia).

Una vez detectada la presencia de una persona por el módulo de visión artificial, el sistema de control se encarga de ejecutar acciones de seguridad a través de un variador de frecuencia (VFD), un PLC Siemens PLC S7-1200 y relés de corte. El objetivo es detener inmediatamente el ascensor y bloquear su funcionamiento mientras se mantenga la presencia humana en la zona de acceso.

Durante las pruebas realizadas con una carga total de 25 kg en la cabina y bajo condiciones de iluminación controladas, se midió el comportamiento eléctrico del sistema. El variador VFD075E23A operó a 9 Hz, manteniendo la velocidad del motor con una desviación inferior al ± 2 % respecto al valor de consigna, lo cual garantiza un control preciso del movimiento del ascensor.

Asimismo, los tiempos de respuesta del PLC y del sistema de relés fueron medidos, mostrando una actuación total inferior a 50 ms, permitiendo cortar el movimiento del motor antes de que la cabina avance más de 2 cm desde el momento de la detección.

La Tabla 3 resume los principales valores eléctricos y tiempos de actuación registrados durante la validación funcional del sistema de control.

TABLA III
DESEMPEÑO ELÉCTRICO Y TIEMPOS DE ACTUACIÓN DEL
SISTEMA DE CONTROL

Parámetro	Valor Medido
Corriente del inversor	2.5 A a 220 V
Estabilidad de velocidad	$\pm 2\%$ a 9 Hz
Tiempo de escaneo del PLC	< 10 ms
Retardo de actuación del relé	< 50 ms
Desplazamiento máximo tras corte	< 2 cm

Estos resultados validan que el sistema de control cumple con los requisitos de respuesta rápida y precisión necesarios para operar como un mecanismo de seguridad efectivo en un entorno industrial simulado.

D. Consumo energético del sistema embebido.

Uno de los objetivos clave en el diseño del sistema fue garantizar su operatividad con un consumo energético bajo, adecuado para entornos donde la eficiencia y la autonomía son relevantes. Para ello, se realizaron mediciones directas de tensión, corriente y potencia consumida por los componentes embebidos durante su funcionamiento normal.

Las pruebas se enfocaron en los dos microcontroladores del sistema: el ESP32-CAM, que ejecuta la inferencia del modelo de visión artificial, y el ESP32-WROOM, encargado de gestionar la lógica de control. Ambos módulos trabajan alimentados a 5V DC, y su consumo fue monitoreado con instrumentos de medición digital.

La Tabla 4 muestra los valores registrados de voltaje, corriente y potencia por cada módulo. Se verificó que el consumo total no superara los 5 W, cumpliendo con el límite establecido en la etapa de diseño del sistema.

TABLA IV
CONSUMO DE ENERGÍA DE LOS MÓDULOS EMBEBIDOS

Componente	Voltaje (V)	Corriente (A)	Potencia (W)
100	5	0.25	1.25
200	5	0.35	1.75
300	5	0.60	3.00

El resultado confirma que el sistema propuesto es energéticamente eficiente, y puede ser alimentado mediante fuentes de bajo consumo, como baterías de litio o paneles solares, sin comprometer su desempeño ni su tiempo de respuesta.

IV. DISCUSIONES

A. Comparación con investigaciones previas.

El sistema propuesto implementa una red SSD-MobileNet en un módulo ESP32-CAM para la detección de personas en entornos industriales, integrando la lógica de control mediante PLC Siemens y VFD. Esta arquitectura se diseñó bajo criterios de eficiencia energética, bajo costo y compatibilidad con infraestructura industrial.

Los resultados obtenidos muestran una precisión del 93.31 % y una F1-score de 93.13 %, superando trabajos similares como el de [15], donde se utilizó YOLOv3 sobre Raspberry Pi para monitoreo en espacios educativos, alcanzando solo un 89 % de precisión bajo condiciones controladas. De forma similar, el estudio de [16] empleó MobileNetV2 en espacios cerrados con una precisión del 90.2 %, pero con un consumo energético superior a 4.5 W, sin integración con sistemas de control industrial.

En el ámbito internacional, se identifican enfoques comparables:

De acuerdo con [17] desarrollaron un sistema basado en TinyML para ascensores sin contacto, alcanzando una precisión de 83.34 % con latencia de 5 segundos, pero con mayor consumo energético que el sistema aquí propuesto.

Por otro lado [18] implementaron MobileNet-SSD sobre Jetson Nano para conteo de personas en elevadores, logrando una precisión de 85 %, aunque su plataforma requiere mayor potencia y consumo eléctrico.

De la misma manera, [19] reportaron que modelos ligeros como SSD-MobileNet pueden reducir su rendimiento hasta 85 % ante condiciones de iluminación variable. En contraste, el presente sistema mantiene una F1-score de 93.13 %, incluso bajo iluminación media y ángulos de visión no ortogonales, lo que evidencia una mayor robustez.

Estas comparaciones permiten concluir que la solución desarrollada:

- Alcanza mayor precisión frente a alternativas en plataformas embebidas y de mayor consumo.
- Optimiza el uso de energía, manteniéndose por debajo de 3 W, lo cual es clave para sistemas autónomos de seguridad.
- Integra exitosamente componentes industriales estándar, como PLC y VFD, sin necesidad de plataformas de alto consumo.

En conjunto, estos resultados consolidan la propuesta como una alternativa robusta, precisa y energéticamente eficiente para mejorar la seguridad en ascensores de carga dentro de entornos industriales reales.

B. Robustez del sistema ante variabilidad operativa

Aunque el sistema fue validado únicamente bajo condiciones de iluminación moderada y en un entorno sin polvo ni vibraciones intensas, los resultados demostraron estabilidad en latencia (240 ms promedio) y baja variabilidad en el frenado del ascensor (≤ 10 cm a 0.4 m/s). La rápida inferencia (~5 fps) y la integración sincrónica con el PLC contribuyeron a una respuesta efectiva ante eventos críticos de seguridad.

Sin embargo, la ausencia de pruebas bajo condiciones extremas (luz intensa, interferencias electromagnéticas, oclusión parcial) representa una limitación para afirmar su desempeño robusto en escenarios industriales reales.

C. Escalabilidad y despliegue real

El diseño modular y el bajo consumo energético (3 W en total) hacen viable su integración en infraestructuras existentes sin necesidad de fuentes de alimentación especializadas. No obstante, su aplicación práctica a ascensores de mayor capacidad requerirá:

- Adaptación del modelo de detección a cámaras de mayor resolución o ángulos múltiples.
- Rediseño del PLC con módulos de entrada/salidas adicionales.
- Incorporación de sistemas redundantes para aumentar la confiabilidad.

D. Limitaciones del estudio

A continuación, se detallan las principales limitaciones del sistema propuesto, con énfasis en sus implicancias para la implementación real y futura escalabilidad.

Validación en entorno restringido. El prototipo fue evaluado en condiciones controladas, utilizando un ascensor de carga de tres niveles y una capacidad máxima de 25 kg. Esta configuración limitada no representa la complejidad de sistemas industriales reales con mayores cargas, más pisos o tráfico continuo. En consecuencia, los resultados deben considerarse preliminares y no generalizables sin pruebas adicionales en entornos reales.

Conjunto de datos limitado. La base de datos utilizada para el entrenamiento del modelo de detección no incluyó imágenes en condiciones extremas, como baja iluminación, contraluz, polvo o borrosidad por movimiento. Aunque se logró una precisión elevada (93.31 %), el modelo podría verse afectado en escenarios más exigentes. Futuras versiones deben incorporar técnicas de aumento de datos y adquisición en campo para mejorar la robustez.

Ausencia de componentes certificados. El sistema emplea relés convencionales y no incluye componentes con certificación de seguridad funcional (por ejemplo, SIL-2). Esto

limita su uso en industrias que operan bajo normativas estrictas de seguridad, especialmente donde se transporten personas o cargas peligrosas. La integración de relés de seguridad e interfaces redundantes sería clave en futuras iteraciones.

Lógica de control básica. El sistema emplea una lógica de antirrebote simple y no utiliza fusión de sensores ni diagnósticos internos. Esto podría afectar su capacidad de respuesta ante condiciones anómalas o fallos del sensor de visión. Para aplicaciones críticas, se recomienda implementar técnicas de detección redundante, watchdogs y registro de eventos.

Costos y mantenibilidad no analizados en profundidad.

Aunque el hardware es de bajo costo (ESP32-CAM), el sistema también incluye un PLC Siemens S7-1200 y un VFD industrial, cuyo costo total puede ser elevado para microempresas. Además, no se aborda la mantenibilidad: limpieza de la lente, actualización del firmware o recalibración periódica. Estos aspectos son esenciales para evaluar la viabilidad operativa a largo plazo.

En resumen, estas limitaciones no invalidan los resultados obtenidos, pero establecen claramente el alcance de este trabajo como una prueba de concepto, y marcan la hoja de ruta para futuras investigaciones orientadas a la validación en entornos industriales complejos y con requisitos normativos.

E. Futuras líneas de mejora

Como trabajos futuros se proponen las siguientes acciones para avanzar hacia una implementación industrial:

- Validar el sistema en entornos operativos reales con diferentes tipos de ascensores.
- Ampliar la base de datos de entrenamiento incluyendo condiciones adversas (iluminación extrema, objetos móviles, polvo).
- Evaluar la robustez del sistema mediante análisis de modos de falla y efectos (FMEA).
- Explorar arquitecturas más ligeras y rápidas como YOLO-Nano o EfficientDet, optimizadas para microcontroladores.
- Escalar el sistema a ascensores con mayor número de niveles y mayores velocidades de operación.

V. CONCLUSIONES

Este artículo presentó el diseño, implementación y validación de un sistema embebido de seguridad para ascensores de carga, basado en visión artificial utilizando la arquitectura SSD-MobileNet sobre un microcontrolador ESP32-CAM, e integrado con un PLC Siemens S7-1200 y un variador de frecuencia (VFD). El sistema logró detectar la presencia de personas con una precisión del 93.31 %, un F1-

score de 93.13 % y un tiempo de respuesta total de 240 ms, lo que permitió detener el ascensor antes de recorrer 0.1 m, cumpliendo así con los requisitos de seguridad industrial.

El consumo energético total del sistema embebido fue inferior a 3 W, lo cual lo hace compatible con fuentes de bajo consumo, como baterías o paneles solares, y adecuado para entornos donde la eficiencia energética es crítica.

A diferencia de soluciones convencionales o implementaciones en plataformas de mayor consumo energético, este trabajo demuestra que es factible integrar visión artificial en tiempo real y lógica de control en dispositivos de bajo costo, manteniendo compatibilidad con infraestructura industrial estándar.

Sin embargo, se identificaron limitaciones importantes: la validación se realizó en un entorno controlado, sin presencia de interferencias electromagnéticas ni condiciones extremas de operación, y con componentes no certificados en seguridad funcional (por ejemplo, relés SIL). Además, el modelo fue entrenado con un conjunto de datos acotado, lo cual limita su generalización a escenarios reales más exigentes.

A pesar de estas limitaciones, el sistema representa una prueba de concepto robusta y viable, que sienta las bases para futuras investigaciones centradas en la validación en campo, el cumplimiento normativo y la integración de arquitecturas más eficientes. Este trabajo constituye un avance relevante hacia la incorporación de inteligencia artificial embebida en aplicaciones industriales críticas de seguridad.

REFERENCES

- [1] Ministerio de Trabajo y Promoción del Empleo (MTPE), "Informe sobre accidentes laborales en plataformas elevadoras," Lima, Perú, 2023.
- [2] F. Cárdenas et al., "Sistema de control para ascensores industriales mediante sensores ópticos," *Rev. Ingeniería y Sociedad*, vol. 12, no. 1, 2021.
- [3] S. Ren et al., "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE TPAMI*, vol. 39, no. 6, pp. 1137–1149, 2017.
- [4] B. Simpson, et al., "Title of paper goes here if known," unpublished.
- [5] W. Liu et al., "SSD: Single Shot MultiBox Detector," in *ECCV*, 2016.
- [6] A. Howard et al., "MobileNets: Efficient CNNs for Mobile Vision," *arXiv:1704.04861*, 2017.
- [7] R. Szeliski, *Computer Vision: Algorithms and Applications*, Springer, 2022.
- [8] J. Muñoz et al., "Implementation of ESP32-CAM for human detection in real-time," *Sensors*, vol. 21, no. 12, 2021.
- [9] Y. Zhao et al., "IoT Visual Monitoring with ESP32-CAM and DL," *MDPI Electronics*, vol. 10, no. 7, 2021.
- [10] G. Flores et al., "Sistema embebido de seguridad con visión artificial," *IEEE Latin America Transactions*, vol. 20, no. 5, 2022.
- [11] K. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," *arXiv:1804.02767*, 2018.
- [12] M. Santana et al., "Aplicación de visión artificial para control de acceso," *Ingenius*, no. 29, pp. 45–53, 2023.
- [13] P. Martínez et al., "Integration of CV and PLCs for Real-Time Industrial Automation," *Springer Industrial Informatics*, vol. 48, 2020.
- [14] R. López et al., "Robustez de visión artificial en entornos industriales," *Revista de Tecnología e Ingeniería*, vol. 25, no. 2, 2022.
- [15] A. Torres et al., "Latency and safety response in embedded vision," *IEEE Access*, vol. 9, pp. 14523–14531, 2021.
- [16] J. Ramírez, J., & V. Vásquez, C. (2022). Sistema de detección de personas para entornos educativos mediante YOLOv3 y Raspberry Pi. Universidad Tecnológica del Perú.
- [17] B. Camacho Chamba, B. (2021). Aplicación de MobileNetV2 para sistemas de vigilancia en interiores. Universidad Técnica del Norte.
- [18] A. Pimpalkar and A. Niture, "Contactless Lift System Using AI Based TinyML for Safety and Automation," *International Journal for Research in Applied Science and Engineering Technology (IJRASET)*, vol. 10, no. 5, pp. 4436–4440, May 2022. doi: 10.22214/ijraset.2022.4265.
- [19] T.-C. Shen and E. T.-H. Chu, "Edge-Computing-Based People-Counting System for Elevators Using MobileNet-Single-Stage Object Detection," *Future Internet*, vol. 15, no. 10, art. 337, Oct. 2023. doi: 10.3390/fi15100337.
- [20] A. Nikouei, S. Wang, Y. Chen, P. Lin, Y. Song, and M. Kantarcioglu, "Real-time Human Detection as an Edge Service Enabled by a Lightweight CNN," in *2018 IEEE International Conference on Edge Computing (EDGE)*, San Francisco, CA, USA, 2018, pp. 125–132. doi: 10.1109/EDGE.2018.00026.