FACULTAD DE INGENIERÍA, ARQUITECTURA Y URBANISMO

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS

TESIS

Redes neuronales convolucionales para la identificación automática de individuos sujetos a búsqueda en terminales terrestres

PARA OPTAR EL TÍTULO PROFESIONAL DE INGENIERO DE SISTEMAS

Autores

Bach. Mondragon Fernandez Alex
ORCID: https://orcid.org/ 0000-0001-8270-392X

Bach. Yarango Farro Darwin Orlando
ORCID: https://orcid.org/ 0000-0002-7704-287X

Asesor

Mg. Mejia Cabrera Heber Ivan
ORCID: https://orcid.org/0000-0002-0007-0928

Línea de Investigación

Ciencias de la información como herramientas multidisciplinares y estratégicas en el contexto industrial y de organizaciones

Sublínea de Investigación

Nuevas tendencias digitales orientadas al análisis y uso estratégico de la información

Pimentel – Perú

2025

DECLARACIÓN JURADA DE ORIGINALIDAD

Quienes suscriben Yarango Farro Darwin Orlando y Mondragon Fernandez Alex, la DECLARACIÓN JURADA, somos Egresados del Programa de Estudios de Ingeniería de Sistemas de la Universidad Señor de Sipán S.A.C, declaramos bajo juramento que somos autores del trabajo titulado:

Redes neuronales convolucionales para la identificación automática de individuos sujetos a búsqueda en terminales terrestres

El texto de mi trabajo de investigación responde y respeta lo indicado en el Código de Ética de la Universidad Señor de Sipán, conforme a los principios y lineamientos detallados en dicho documento, en relación con las citas y referencias bibliográficas, respetando el derecho de propiedad intelectual, por lo cual informo que la investigación cumple con ser inédita, original y auténtica.

A su vez Declaro bajo juramento que todas las firmas consignadas en los documentos adjuntos a esta investigación son auténticas y han sido obtenidas con el consentimiento de los firmantes. En caso de detectarse falsificación, asumo plena responsabilidad académica y administrativa, sujetándome a las sanciones establecidas en el Código de Ética y Reglamento de la Universidad Señor de Sipán.

En virtud de lo antes mencionado, firman:

Mondragon Fernandez Alex	DNI: 60160683	
Yarango Farro Darwin Orlando	DNI: 76680636	

Pimentel, 10 de mayo del 2025

REPORTE DE SIMILITUD TURINITIN

Mondragon Fernandez Alex Yarango Farro Darwin ...

Redes neuronales convolucionales para la identificación automática de individuos sujetos a búsqueda

Universidad Señor de Sipan

Detalles del documento

Identificador de la entrega
trn:oid:::26396:469965129

Fecha de entrega
26 jun 2025, 12:40 p.m. GMT-5

Fecha de descarga
26 jun 2025, 12:41 p.m. GMT-5

Nombre de archivo
Turnitin NFORME_FINAL-YARANGO_MONDRAAGON.docx

Tamaño de archivo
2.8 MB

24 Páginas

4728 Palabras

27.328 Caracteres

turnitin Página 2 of 29 - Descripción general de integridad Identificador de la entrega trn:oid:::26396:469965129

12% Similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para ca...

Filtrado desde el informe

- Bibliografía
- Texto mencionado
- Coincidencias menores (menos de 8 palabras)

Fuentes principales

9% Fuentes de Internet
0% Publicaciones
4% Trabajos entregados (trabajos del estudiante)

Marcas de integridad

N.º de alertas de integridad para revisión

No se han detectado manipulaciones de texto sospechosas.

Los algoritmos de nuestro sistema analizan un documento en profundidad para buscar inconsistencias que permitirían distinguirlo de una entrega normal. Si advertimos algo extraño, lo marcamos como una alerta para que pueda revisarlo.

Una marca de alerta no es necesariamente un indicador de problemas. Sin embargo, recomendamos que preste atención y la revise.

REDES NEURONALES CONVOLUCIONALES PARA LA IDENTIFICACIÓN AUTOMÁTICA DE INDIVIDUOS SUJETOS A BÚSQUEDA EN TERMINALES TERRESTRES

Aprobación del jurado

DR. OLIVER VASQUEZ LEYVA
Presidente del Jurado de Tesis

MG. JAIME ARTURO BRAVO RUIZ
Secretario del Jurado de Tesis

MG. HEBER IVAN MEJIA CABRERA
Vocal del Jurado de Tesis

REDES NEURONALES CONVOLUCIONALES PARA LA IDENTIFICACIÓN AUTOMÁTICA DE INDIVIDUOS SUJETOS A BÚSQUEDA EN TERMINALES TERRESTRES

Resumen

El objetivo de este estudio fue desarrollar un sistema automático de identificación de individuos buscados en terminales terrestres utilizando Redes Neuronales Convolucionales (CNN). Se realizó bajo un enfoque cuantitativo y un diseño cuasi-experimental. Se empleó un conjunto de datos privado con fines educativos, compuesto por imágenes y videos de individuos en un entorno dinámico, lo que permitió pruebas en contextos reales. La metodología incluyó la carga estructurada de datos biométricos, la detección facial mediante el modelo Multi-Task Cascaded Convolutional Neural Network (MTCNN) y la generación de embeddings faciales con el modelo InceptionResNetV1. Las características extraídas se almacenaron en una base de datos MySQL. Para optimizar la búsqueda en tiempo real, los embeddings se transfirieron a FAISS, una biblioteca optimizada para la búsqueda de similitudes en grandes volúmenes de datos, donde se almacenaron en formato vectorial para una búsqueda eficiente. Luego, se realizó la identificación en secuencias de video en tiempo real. Los resultados mostraron un alto desempeño del sistema, con una precisión del 89% en la detección de rostros y del 98.60% en la identificación de personas, al compararlas con los datos almacenados. Finalmente, los modelos entrenados se integraron en una aplicación web que permite la búsqueda mediante cámaras IP, utilizando la arquitectura YOLO versión 8 para el seguimiento de personas identificadas. Estos resultados confirman que los modelos de deep learning pueden incorporarse eficazmente en sistemas de vigilancia y control en espacios públicos, mejorando la seguridad y ofreciendo una solución viable para la identificación automática en tiempo real.

Palabras clave: Redes Neuronales Convolucionales, identificación automática, reconocimiento facial, vigilancia en tiempo real, terminales terrestres.

Abstract

The objective of this study was to develop an automatic system for identifying wanted individuals in terrestrial terminals using Convolutional Neural Networks (CNN). It was conducted under a quantitative approach and a quasi-experimental design. A private dataset intended for educational purposes was used, consisting of images and videos of individuals in dynamic environments, enabling testing in real-world contexts. The methodology included the structured loading of biometric data, facial detection using the Multi-Task Cascaded Convolutional Neural Network (MTCNN), and the generation of facial embeddings with the InceptionResNetV1 model. The extracted features were stored in a MySQL database. To optimize real-time search, the embeddings were transferred to FAISS, a library optimized for similarity search in large-scale datasets, where they were stored in vector format to enable fast and efficient retrieval. Subsequently, identification was performed on video sequences captured in real time. The results demonstrated high system performance, with an accuracy of 89% in face detection and 98.60% in person identification when compared with stored data. Finally, the trained models were integrated into a web application that enables real-time search via IP cameras, using the YOLO version 8 architecture for tracking identified individuals. These findings confirm that deep learning models can be effectively incorporated into surveillance and control systems in public spaces, enhancing existing security mechanisms and providing a viable solution for automatic real-time person identification.

Keywords: Convolutional Neural Networks, automatic identification, facial recognition, real-time surveillance, terrestrial terminals.

I. INTRODUCCIÓN

En la era digital actual, la identificación automática de individuos se ha convertido en un aspecto crucial en diversas aplicaciones, desde la seguridad pública hasta la gestión de accesos en instalaciones sensibles. La necesidad de sistemas de identificación precisos y eficientes ha llevado al desarrollo de tecnologías avanzadas que pueden procesar y analizar datos visuales en tiempo real [1].

En este contexto, la inseguridad pública constituye un desafío significativo a nivel global, adquiriendo una relevancia particular en los terminales de transporte terrestre, los cuales, además de ser nodos estratégicos en la movilidad de personas y mercancías, se presentan como áreas altamente vulnerables a actividades delictivas transnacionales. Estos espacios han sido identificados como puntos críticos en los esfuerzos internacionales para mitigar el crimen organizado, lo que subraya la necesidad imperiosa de reforzar la cooperación internacional y de modernizar los sistemas de seguridad para enfrentar de manera efectiva las amenazas emergentes [2].

En Perú, la situación no es muy diferente. Los métodos de identificación manual empleados en los terminales terrestres han ralentizado los operativos de seguridad, afectando la capacidad de las autoridades para actuar con rapidez y precisión. Según el Instituto Nacional de Estadística e Informática (INEI), el sistema de verificación del Documento Nacional de Identidad (DNI), utilizado para identificar a posibles prófugos de la justicia, es un proceso lento y frecuentemente ineficaz [3]. Esta lentitud no solo retrasa las operaciones policiales, sino que también incrementa el riesgo de que personas con órdenes de captura eludan la vigilancia. Asimismo, en Cusco, las autoridades continúan utilizando métodos tradicionales como la exhibición de carteles con fotografías de los “más buscados”, una técnica que, aunque útil en ciertos contextos, resulta obsoleta frente a la creciente sofisticación de las actividades delictivas [4].

Esta situación se agudiza particularmente en ciudades como Chiclayo, donde el problema resulta aún más acuciante, durante el año 2023, la oficina de la Defensoría del Pueblo en Chiclayo emitió un informe en el que alertó sobre el preocupante incremento de delitos contra el patrimonio, como robos, hurtos y asaltos a mano armada, en los distritos de La Victoria y José Leonardo Ortiz, considerados los más vulnerables en materia de seguridad ciudadana [5]. Según datos proporcionados por la Policía Nacional del Perú, solo en el primer semestre del año se reportaron más de 2,800 denuncias por hechos delictivos en la provincia de Chiclayo, de los cuales un porcentaje significativo ocurrió en zonas bajo vigilancia municipal [6].

Pese a contar con una red aproximada de 450 cámaras de videovigilancia distribuidas en espacios públicos estratégicos, se constató que más del 65 % de estos dispositivos presentaban fallas técnicas, resoluciones inadecuadas para el reconocimiento facial, o bien estaban desconectadas de los centros de control. Además, la mayoría de estos equipos carece de integración con sistemas de análisis automático de imágenes, lo que impide realizar tareas de identificación facial en tiempo real. Este hecho fue señalado como una grave deficiencia, ya que limita la capacidad de las autoridades para actuar con rapidez ante la presencia de personas con requisitorias judiciales o involucradas en actividades delictivas [7].

Actualmente, el monitoreo de las cámaras se realiza de forma manual por operadores humanos, quienes deben observar constantemente las transmisiones en directo sin apoyo de algoritmos que faciliten la detección de rostros o comportamientos sospechosos. Esto no solo genera una sobrecarga en el personal, sino que también incrementa el margen de error y reduce la efectividad del sistema de videovigilancia.

Esta carencia tecnológica no solo dificulta la identificación de individuos buscados, sino que también incrementa la percepción de inseguridad en la población. Para abordar estos problemas, se requiere con urgencia la implementación de soluciones tecnológicas avanzadas, como las redes neuronales convolucionales (CNN).

Modelos como FaceNet, diseñados específicamente para el análisis de imágenes y reconocimiento de patrones, han demostrado una precisión del 97% en pruebas controladas de identificación facial, lo que los convierte en herramientas idóneas para reforzar los sistemas de vigilancia y control en los terminales terrestres [8].

Las CNN han emergido como una de las herramientas más efectivas en el campo del aprendizaje profundo, especialmente en tareas de reconocimiento de patrones y clasificación de imágenes. Su capacidad para aprender características jerárquicas a partir de datos visuales las hace particularmente adecuadas para la identificación de individuos a través de imágenes faciales y otros datos biométricos [9][10].

En el contexto de los terminales terrestres, las CNN pueden ser implementadas para mejorar la seguridad y la eficiencia en la identificación de personas. Estas redes no solo permiten la identificación precisa en condiciones variadas, sino que también pueden integrarse en sistemas de vigilancia y control de acceso, facilitando una respuesta rápida ante situaciones de seguridad [11] [12] [13].

Estudios recientes han explorado el potencial de las tecnologías avanzadas para optimizar la vigilancia y la identificación de personas buscadas.

En India, el uso de CNN como VGG-16 ha mostrado resultados prometedores en la identificación de personas desaparecidas mediante técnicas avanzadas como el preprocesamiento de datos y el aumento de muestras, se logró una precisión del 90.01% durante el entrenamiento y del 85.21% en las pruebas; estos resultados validan el potencial de las CNN para abordar problemas complejos de localización e identificación en entornos terrestres [14].

Las CNN han sido utilizadas en la Reidentificación de personas (ReID) en sistemas de videovigilancia, en el trabajo [15] se analizó las arquitecturas de CNN para la reidentificación de personas en sistemas de videovigilancia, evaluó la efectividad de diferentes sistemas utilizando diversos conjuntos de datos, destacando la importancia de la extracción de características y los métodos de entrenamiento. En [16] con el objetivo de mejorar la precisión de la ReID en imágenes de sistemas de

videovigilancia distribuidos, esta investigación utilizó conjuntos de datos como Market1501 y PolReID, se investigaron diversas funciones de activación (ReLU, Leaky-ReLU, PReLU) y arquitecturas de CNN (ResNet-50, DenseNet-121, DarkNet-53), mostrando mejoras significativas en las métricas de precisión.

Las CNN también se están utilizando en el conteo de multitudes en la gestión urbana, en el trabajo [17] en el contexto de la gestión urbana y la seguridad pública, se utilizó un modelo pequeño de CNN para el conteo de multitudes en terminales embebidos, el modelo demostró ser efectivo, eficiente y adaptable, logrando un equilibrio entre precisión y eficiencia en las aplicaciones de conteo de multitudes.

Por su parte, en Huancayo, la Policía Nacional implementó un sistema basado en redes neuronales convolucionales para identificar puntos faciales únicos, fortaleciendo así el control en zonas estratégicas de vigilancia [18].

Desde el punto de vista teórico, las redes neuronales convolucionales, basadas en los principios del aprendizaje profundo propuestos por Geoffrey Hinton [19], han revolucionado el análisis de imágenes y el reconocimiento de patrones faciales complejos. Estas redes, mediante múltiples capas de procesamiento, son capaces de extraer y reconocer características faciales con un nivel de precisión sin precedentes [20]. Además, la técnica de transferencia de aprendizaje permite utilizar modelos previamente entrenados en grandes bases de datos y adaptarlos a nuevas tareas, como la identificación de personas buscadas, reduciendo de manera significativa el tiempo y los recursos necesarios para implementar soluciones efectivas de reconocimiento facial en diversos entornos [20]. Además, la técnica de transferencia de aprendizaje permite utilizar modelos previamente entrenados en grandes bases de datos y adaptarlos a nuevas tareas, como la identificación de personas buscadas [21]. Esto reduce de manera significativa el tiempo y los recursos necesarios para implementar soluciones efectivas de reconocimiento facial en diversos entornos [22].

Estos estudios demuestran colectivamente la versatilidad y efectividad de las CNN

en la mejora de la identificación de personas, desde la videovigilancia y el reconocimiento facial hasta la gestión de multitudes.

El objetivo del presente estudio fue desarrollar un sistema de identificación automática basado en redes neuronales convolucionales para su implementación en los terminales de transporte terrestres de Chiclayo, ciudad al norte del Perú. Este sistema busca mejorar la seguridad pública mediante la identificación rápida y precisa de individuos como una herramienta de búsqueda para el apoyo en la gestión de las fuerzas de seguridad.

II. MATERIALES Y MÉTODO

2.1 Materiales

1. Conjunto de datos de imágenes

En esta investigación se utilizó un conjunto de personas en condición de búsqueda, compuesto por 700 imágenes en formato PNG con dimensiones de 600 × 840 píxeles y secuencias de video de individuos capturados en un entorno dinámico, lo que permitió realizar pruebas de identificación en contextos reales, con variaciones propias del flujo natural de personas en terminales terrestres. En la Figura 1 se muestran dos imágenes de muestra extraídas del conjunto de datos, correspondientes a los autores del estudio, que reflejan tanto la calidad de resolución como las condiciones ambientales bajo las cuales se entrenó y evaluó el sistema propuesto.

Figura 1. Imágenes de muestra del dataset

2. Herramientas de software

En la Tabla 1 se enumeran las herramientas utilizadas para la identificación de individuos sujetos a búsqueda en terminales terrestres.

Tabla 1. Sofware utilizado para el entrenamiento

Tipo	Nombre	Licencia de Uso	Descripción
Etiquetación	Roboflow	Comercial	Herramienta para etiquetar y gestionar datos de imágenes.
Arquitectura CNN	MTCNN	MIT	Red CNN para detectar rostros con precisión.
Arquitectura CNN	FastNet	MIT	Red CNN ligera para comparación de rostros.
Arquitectura CNN	Yolo v8	Open Source	Modelo rápido para seguimiento de individuos en tiempo real.
Lenguaje de Programación	Python	Python License	Lenguaje usado en desarrollo de modelos de IA.

Fuente: Elaboración del autor

En la Tabla 2 se enumeran las herramientas necesarias para desarrollar la aplicación web que permitió realizar la identificación automática de individuos sujetos a búsqueda.

Tabla 2. Sofware para la aplicación web

Tipo	Nombre	Licencia de Uso	Descripción
Lenguaje de programación	HTML, CSS, JavaScript	MIT / Open Source	Desarrollo visual e interacción en el navegador.
Lenguaje de programación	Python	Python License	Lógica y funcionamiento del servidor.

Fuente: Elaboración del autor

3. Servicio

Para el entrenamiento del modelo de identificación se ha utilizado el servicio de Google Collaboratory cuyas características se enumeran en la tabla 3.

Tabla 3. Entorno de entrenamiento de los modelos.

Nombre	Característica	Valor
Google Collaboratory	Velocidad	8 CPUs virtuales a 3.7 GHz
	Almacenamiento	107.7 Gb
GPU NVIDIA T4	Memoria RAM	12,7 Gb
	Memoria GPU	15 Gb GDDR6

Fuente: Elaboración del autor

4. Hardware

En la tabla 4, se describen las características del equipo donde se ha desplegado la aplicación web.

Tabla 4. Características del servidor para el despliegue

Nombre	Característica	Valor
Servidor Web	Velocidad del procesador	2 CPU 2,50 GHz
	Almacenamiento	10 Gb
	Memoria RAM	8 Gb
	Servidor Web	Flask (Python 3.10)

Fuente: Elaboración del autor

2.2 Métodos

Esta investigación tecnológica aplicada, que emplea un enfoque cuantitativo para recopilar y analizar datos de forma objetiva [23] [24], se centra en abordar un problema práctico mediante la implementación de CNN para la identificación automática de individuos bajo orden de búsqueda en terminales terrestres. El estudio adopta un

diseño cuasi-experimental, que se caracteriza por la ausencia de selección aleatoria para los grupos experimentales, trabajando en su lugar con grupos preexistentes [25]. Inicialmente se ha creado una base de datos que almacena información de las personas buscadas, incluyendo fotografías de sus rostros y los vectores de características correspondientes. El método propuesto permite la búsqueda en video en tiempo real para la detección e identificación de dichos individuos en los terminales terrestres.

1. Detección y preparación de datos faciales

El procedimiento para procesar las imágenes faciales se compone de las siguientes etapas:

- **Carga de datos**

  El procedimiento comienza con la carga de imágenes faciales desde un directorio predefinido. Cada archivo de imagen en formato PNG representa a un individuo, y el nombre del archivo contiene su número de documento de identidad (DNI), lo cual permite mantener una vinculación estructurada entre los datos biométricos y la identidad correspondiente, además se ingresa los datos personales.

- **Detección de rostros**

  El procedimiento incluye una validación inicial de rostros, utilizando el modelo MTCNN para detectar rostros en las imágenes, asegurando que solo se procesen regiones faciales válidas [26] .

- **Aumento de datos**

  Una vez cargada una imagen, se aplican 5 transformaciones para aumentar la diversidad de representaciones faciales. Estas incluyen ajustes de brillo, contraste y nitidez, además de rotaciones leves y reflejos horizontales. Tales modificaciones introducen variabilidad en las condiciones de captura, mejorando la capacidad de generalización del sistema de reconocimiento facial.

- **Generar embedding**

  Tanto la imagen original como sus versiones aumentadas son procesadas con InceptionResNet, un modelo de aprendizaje profundo especializado en análisis facial, que identifica con precisión la región del rostro en cada imagen. Si no se detecta ningún rostro, la imagen se descarta del flujo de procesamiento.

  A partir del rostro detectado, se genera un vector de características o embedding facial, que representa matemáticamente las particularidades del rostro en un espacio multidimensional y constituye la unidad básica para tareas de comparación y reconocimiento.

- **Almacenamiento**

  Finalmente, cada embedding se almacena en una base de datos relacional, asociado al individuo correspondiente. La base está estructurada en dos tablas principales: una para los datos personales y otra para los embeddings faciales. Cada registro en la tabla de embeddings mantiene una clave foránea con la tabla de personas, lo que permite almacenar múltiples vectores biométricos por individuo sin duplicar información. Esta estrategia garantiza la integridad referencial y facilita consultas eficientes sobre el historial biométrico de cada persona registrada.

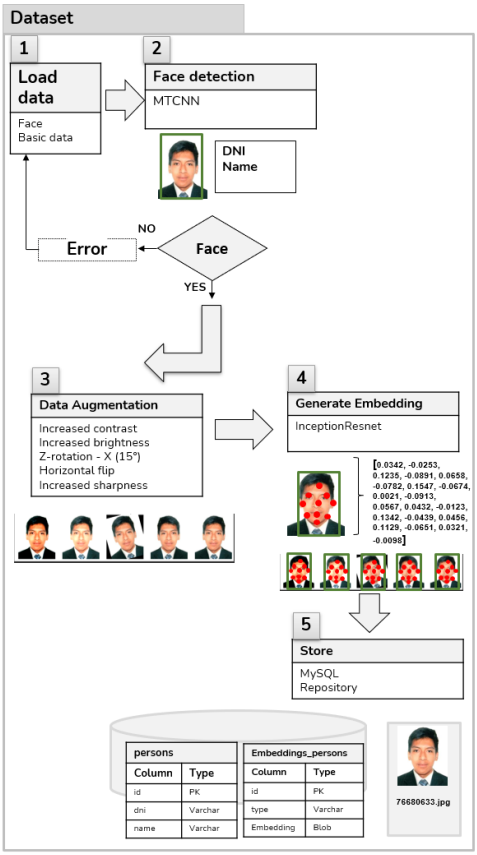

Figura 2. Generación de Embeddings y Almacenamiento para Reconocimiento Facial.

Fuente: Elaboración del autor

- **Datos temporales**

El proceso de carga temporal combina MySQL y FAISS para gestionar información estructurada y facilitar la búsqueda eficiente de *embeddings*. Inicialmente, los datos personales se almacenan en MySQL en dos tablas: Persons, con identificadores, nombres y documentos, y Embeddings_persons, donde se asocian *embeddings* en formato BLOB. Luego, estos *embeddings* se transfieren a FAISS, optimizado para búsquedas en espacios de alta dimensión, lo que permite consultas rápidas y eficientes. Esta integración combina la fiabilidad de MySQL con el rendimiento de un índice vectorial, optimizando la gestión y recuperación de datos en aplicaciones de similitud a gran escala. La figura 3 muestra este proceso.

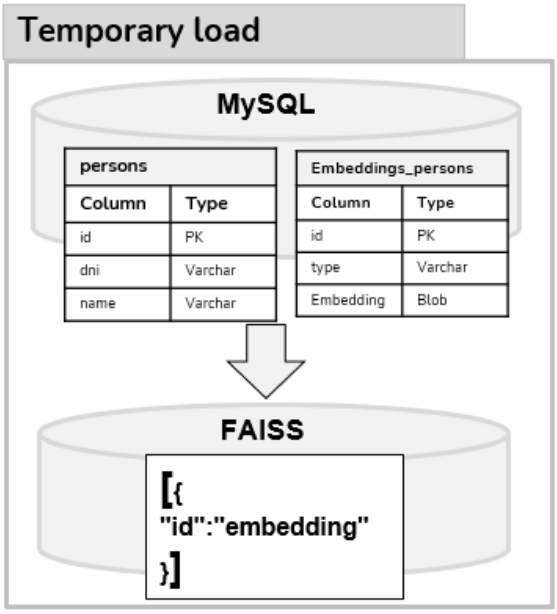

Figura 3. Almacenamiento temporal de datos relacionales en FAISS.

Fuente: Elaboración del autor

2. Identificación y seguimiento

- **Identificación**

La fase de identificación de personas en video inicia con la captura de fotogramas desde una cámara IP, seguida de la detección de rostros utilizando el modelo MTCNN. Posteriormente, el modelo InceptionResNet, aplicado junto con la técnica faceNet, genera un embedding para cada rostro, convirtiéndolo en un vector numérico único que representa sus características faciales.

- **Buscar**

Una vez generado el embedding, se utiliza FAISS [27], para comparar dicho vector con los registros almacenados en la base de datos de embeddings, con el fin de determinar si el rostro corresponde a una persona previamente registrada.

- **Seguimiento**

Tras la identificación, el sistema activa la fase de seguimiento utilizando el modelo YOLOv8 [28], para el tracking en video. Este modelo de detección de objetos permite rastrear a la persona en tiempo real a medida que se desplaza dentro de la escena, garantizando un monitoreo continuo y preciso

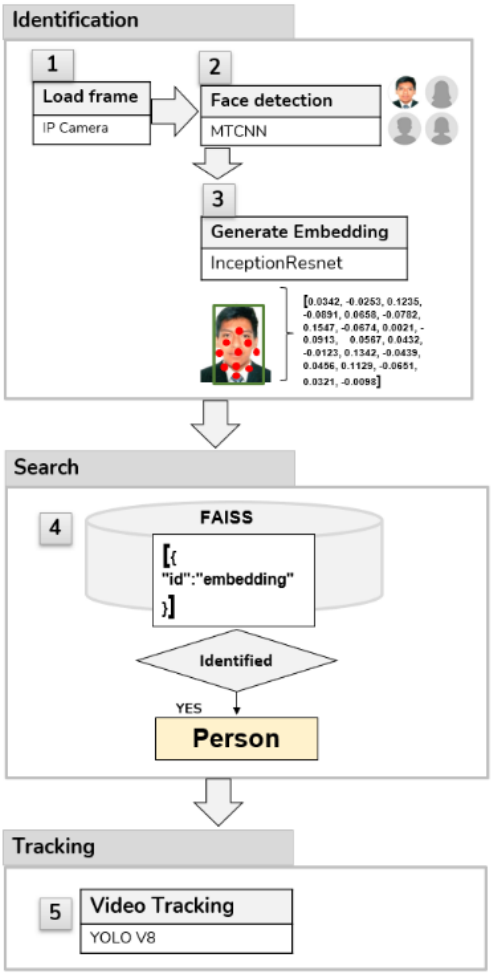

Figura 4. Flujo de seguimiento e identificación facial en tiempo real.

Fuente: Elaboración del autor

3. Aplicación Web

La Figura 5 muestra la arquitectura general de la aplicación web, basada en un enfoque cliente-servidor. La capa de presentación ha sido desarrollada con JavaScript, utilizando HTML y CSS para estructurar y diseñar el entorno gráfico que interactúa con el usuario, esta interfaz permite al usuario enviar solicitudes al servidor, ingresar datos y cargar imágenes de personas que serán procesadas posteriormente.

La comunicación entre la interfaz y el servidor se realiza mediante peticiones HTTP, enviando la información en formato JSON (JavaScript Object Notation), lo que asegura una transferencia eficiente y estructurada de los datos.

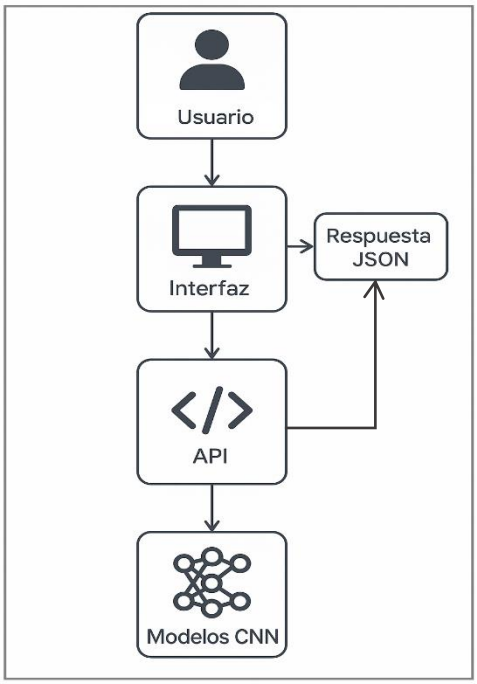

Figura 5. Arquitectura del sofware

Fuente: Elaboración del autor

4. Evaluación del desempeño

Para evaluar el rendimiento del sistema de reconocimiento facial propuesto en la identificación de individuos en terminales terrestres, las predicciones se clasifican según su precisión. Se registra un Verdadero Positivo (VP) cuando el sistema identifica correctamente a una persona que está en la base de datos de buscados. Un Verdadero Negativo (VN) ocurre cuando determina correctamente que una persona no buscada no está en la base de datos. Un Falso Positivo (FP) se da cuando identifica erróneamente a una persona inocente como buscada, y un Falso Negativo (FN) cuando no reconoce a una persona que sí está en la lista. Se utilizaron las siguientes métricas para evaluar el desempeño del sistema:

Precisión (P): Es una métrica que cuantifica la proporción de verdaderos positivos con respecto al número total de elementos clasificados como positivos, evaluando así la exactitud de las predicciones positivas del modelo.

$$P = \frac{TP}{TP + FP} \qquad (1)$$

Recall (R): Evalúa la capacidad del modelo para identificar correctamente todos los casos positivos reales.

$$R = \frac{TP}{TP + FN} \qquad (2)$$

Accuracy (A): Proporción de predicciones correctas realizadas por el modelo en relación con el número total de predicciones efectuadas.

$$A = \frac{TP + TN}{TP + TN + FP + FN} \qquad (3)$$

F1-score (F): Proporciona una medida equilibrada entre la precisión y la sensibilidad (recall).

$$F = 2 * \frac{P * R}{P + R} \qquad (4)$$

III. RESULTADOS Y DISCUSIÓN

3.1 Resultados

El entrenamiento fue realizado en un entorno de computación en la nube, cuyas especificaciones técnicas se presentan en la Tabla 5. Esta configuración de hardware proporcionó los recursos necesarios para la ejecución de los modelos y permitió analizar su rendimiento de manera consistente.

Tabla 5. Configuración del entorno de ejecución experimental

Valor	Métrica
Procesador	Virtualizado CPU
Memoria	12.7 GB RAM
Sistema Operativo	Linux-based (Ubuntu)
Almacenamiento	107.7 GB

Fuente: Elaboración del autor

1. Rendimiento del modelo MTCNN – Detección de rostros

En la Figura 6 se exploran los fotogramas de un video, donde los rostros fueron etiquetados manualmente para indicar sus ubicaciones reales y compararlos con las detecciones realizadas por el modelo MTCNN, este modelo detecta la presencia de rostros y verifica si estas detecciones coinciden con las ubicaciones reales en los fotogramas del video (Figura 7).

Figura 6. Etiquetación manual

Fuente: [29]

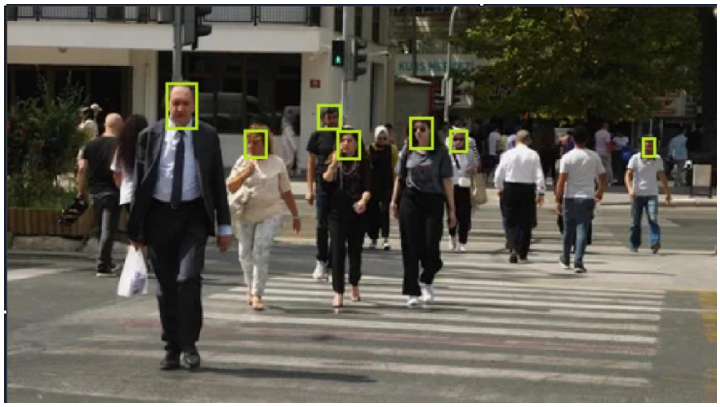

Figura 7. Detección con el modelo MTCNN

Fuente: [29]

En la Tabla 6 se presentan las métricas de evaluación del modelo MTCNN, obtenidas a partir del procesamiento de un video compuesto por 150 fotogramas. Para la validación, se utilizó un umbral de Intersección sobre la Unión (IoU) de 0.3, comparando las predicciones generadas por el modelo con las anotaciones manuales. Los resultados reflejan un buen desempeño en la detección de personas durante la etapa de entrenamiento, alcanzando una precisión del 89 %, un Recall del 78 %, una exactitud del 93 % y un F1-score del 82 %.

Tabla 6. Métricas de rendimiento del modelo MTCNN

Métrica	Valor
Precision	89%
Recall	78%
Accuracy	93%
F1 - score	82%

Fuente: Elaboración del autor

Así mismo la matriz de confusión (Figura 8), proporciona una representación visual del rendimiento del modelo en la detección de rostros de personas en la fase de pruebas, el modelo identificó correctamente 150 casos en los que había una persona presente y 120 casos en los que no había personas prediciendo "No Persona" de forma adecuada,

sin embargo, cometió 2 errores al predecir que no era una persona cuando en realidad lo era y otros 2 errores al predecir que era una persona cuando no lo era.

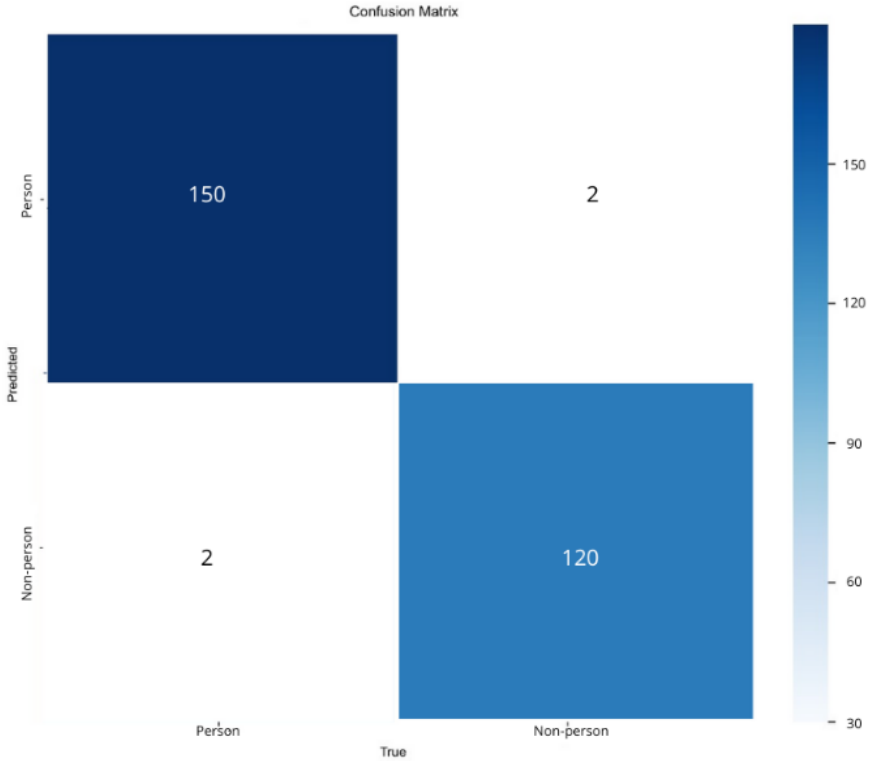

Figura 8. Matriz de confusión que muestra el rendimiento de MTCNN

Fuente: Elaboración del autor

MTCNN se encarga de detectar los rostros en la imagen o video, recortando las regiones correspondientes. Luego, estas regiones de rostro se pasan al modelo InceptionResNetV1, que extrae las características faciales únicas, también conocidas como embeddings, que representan la "huella" facial de cada persona.

2. Rendimiento del modelo FaceNet – Identificación de Personas

En la Tabla 7 se presentan las métricas de evaluación obtenidas durante el proceso de identificación facial. Debido a que se disponía de menos de tres imágenes por persona, fue necesario aplicar técnicas de aumento de datos, como rotación, cambio de escala y ajustes de brillo y contraste, con el objetivo de incrementar la cantidad y diversidad de datos disponibles por individuo. Estas transformaciones permitieron generar

embeddings representativos utilizando el modelo InceptionResNetV1. Los resultados obtenidos reflejan un buen desempeño del modelo FaceNet en la etapa de entrenamiento, alcanzando una precisión del 98.60 %, un Recall del 97.70 %, una exactitud del 99.20 % y un F1-score del 98.15 %.

Tabla 7. Métricas de rendimiento del modelo FaceNet

Métrica	Valor
Precisión	98.60 %
Recall	97.70 %
F1 - score	98.15%
Accuracy	99.20 %

Fuente: Elaboración del autor

Finalmente, se almaceno estos embeddings en una base de datos y se utilizó junto con FaceNet y FAISS para verificar identidades en tiempo real.

3. Rendimiento del modelo YOLO V8 – Seguimiento de Personas

El modelo YOLOv8 asigna un identificador único (ID) a cada persona que detecta, lo que permite hacer un seguimiento en tiempo real, incluso si la persona se mueve o se cruza con otras.

Según las métricas mostradas en la Figura 9, el modelo tiene una alta precisión, lo que significa que acierta en la mayoría de sus detecciones, y también ofrece una buena estimación de la ubicación de las personas (MOTP).

Sin embargo, el recall y el F1-Score son bajos, lo que indica que el modelo puede pasar por alto algunas detecciones. A pesar de eso, el sistema mantiene un rendimiento aceptable en general (MOTA) y es capaz de seguir correctamente a cada individuo, aunque haya oclusiones o cruces entre personas.

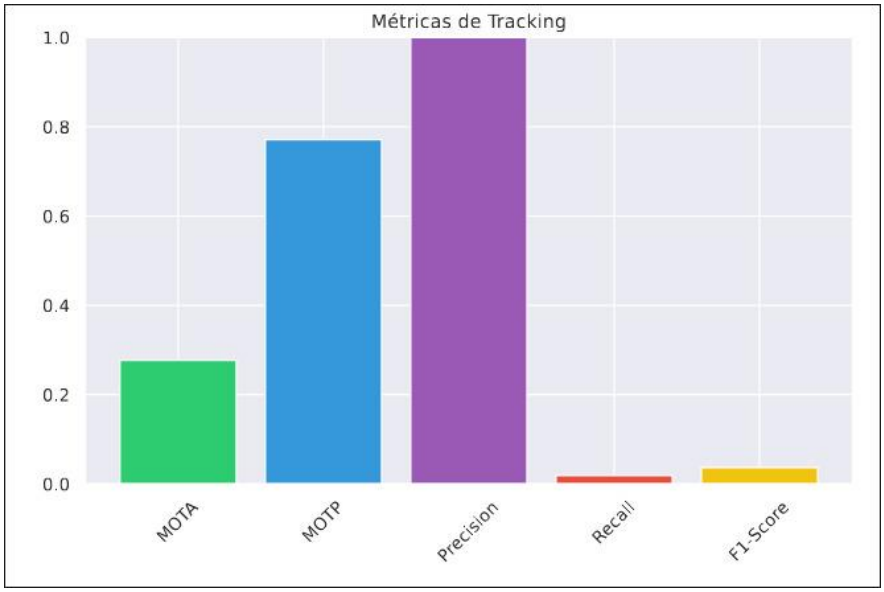

Figura 9. Rendimiento del modelo YOLO v8

Fuente: Elaboración del autor

YOLOv8 demostró ser una herramienta eficaz para el seguimiento de personas en tiempo real, destacando por su alta precisión y estabilidad en la identificación, aunque con áreas de mejora en la detección completa de todos los individuos presentes en la escena.

3.2 Discusión

Los resultados obtenidos mostraron un desempeño robusto del sistema propuesto, compuesto por MTCNN para la detección de rostros, FaceNet con InceptionResNetV1 para la identificación de personas y YOLO versión 8 para el seguimiento en tiempo real. El sistema se diseñó para funcionar eficazmente bajo condiciones de trabajo reales, utilizando un entorno computacional moderado y conjuntos de datos con un número limitado de imágenes por individuo. A pesar de estas restricciones, se alcanzaron resultados destacados que refuerzan la aplicabilidad de este enfoque en escenarios de monitoreo y seguridad.

En la etapa de detección facial, el modelo MTCNN logró una precisión del 89%, un *recall* del 78% y un F1-score del 82%, estos resultados indicaron una alta capacidad del modelo para localizar rostros en condiciones desafiantes de variabilidad de

iluminación y poses. El rendimiento obtenido fue comparable al de enfoques previos como el VGG16, comúnmente utilizado para tareas de reconocimiento facial en entornos con variaciones controladas [30]. Sin embargo, la ventaja del MTCNN radicó en su habilidad para realizar una detección más precisa y eficiente en escenas más dinámicas, gracias a su arquitectura jerárquica que optimiza la localización facial multiescala.

En cuanto a la identificación de personas, el modelo FaceNet, respaldado por InceptionResNetV1, presentó métricas sobresalientes. Con solo tres imágenes por persona en el conjunto de datos, se aplicaron técnicas avanzadas de *data augmentation*, como rotación, escalado y modificaciones de brillo y contraste, lo que permitió expandir efectivamente el conjunto de entrenamiento. Como resultado, se obtuvieron una precisión del 98.60%, un *recall* del 97.70% y una exactitud del 99.20%, valores que colocan al modelo entre los más eficientes en su categoría. En comparación con enfoques más recientes que integran mecanismos de atención multiescala para gestionar oclusiones y variabilidad extrema [13], este estudio demuestra que una adecuada estrategia de *data augmentation* y la selección de una arquitectura profunda adecuada pueden competir efectivamente en cuanto a rendimiento, sin necesidad de incrementar excesivamente la complejidad computacional.

La arquitectura YOLO versión 8, utilizado para el seguimiento de personas en tiempo real, mostró un comportamiento prometedor, aunque aún no se dispone de métricas definitivas. Las primeras pruebas sugirieron que el modelo puede mantener un seguimiento continuo y coherente de identidades en secuencias de video, con una baja tasa de errores en la asignación de identidad y la detección de objetos. Esto lo posiciona como una opción valiosa para aplicaciones de vigilancia y control en tiempo real, superando a los modelos tradicionales orientados al conteo de multitudes, los cuales ofrecen estimaciones globales, pero no pueden realizar un seguimiento individualizado eficaz de personas en escenarios dinámicos [15].

En términos generales, el sistema demostró una integración efectiva de los componentes seleccionados, equilibrando precisión, eficiencia y escalabilidad. No obstante, se identificaron áreas de mejora que podrían fortalecer aún más el rendimiento del sistema. En particular, el número limitado de imágenes por individuo, aunque mitigado mediante *data augmentation*, restringió la capacidad del modelo para generalizar en condiciones de alta variabilidad. Además, el umbral IoU de 0.3 empleado para la evaluación de la detección, si bien aceptable para pruebas iniciales, podría optimizarse para aplicaciones de alta precisión, especialmente en escenarios con más variabilidad o condiciones difíciles de iluminación.

El sistema desarrollado no solo demostró ser efectivo para la detección, identificación y seguimiento de personas, sino que también mostró un excelente equilibrio entre precisión y eficiencia computacional, este trabajo sienta las bases para futuras implementaciones en entornos de monitoreo en tiempo real, control de accesos y aplicaciones de seguridad pública, con el potencial de adaptarse a una gama más amplia de escenarios y condiciones operativas. La escalabilidad del sistema también sugiere su viabilidad para su implementación en sistemas de vigilancia masiva, abriendo nuevas avenidas de investigación para mejorar la robustez en contextos más complejos.

IV. CONCLUSIONES Y RECOMENDACIONES

4.1 Conclusiones

La presente investigación demostró que la aplicación de CNN permite desarrollar un sistema eficaz para la identificación automática de individuos buscados en terminales terrestres. Mediante un enfoque cuantitativo y un diseño cuasi-experimental, se implementó un procedimiento integral que abarca la carga estructurada de datos, la detección y el procesamiento de imágenes faciales, cuya efectividad fue validada empleando un conjunto de datos privado. En este sentido, se logró elaborar un dataset específico de

personas en condición de búsqueda, conformado por imágenes faciales debidamente etiquetadas y organizadas, lo cual resultó esencial para el entrenamiento y evaluación de los modelos implementados. La calidad y estructura de este conjunto de datos contribuyeron significativamente al desempeño general del sistema, evidenciando la importancia de contar con bases de datos representativas y bien curadas para aplicaciones de reconocimiento facial.

El sistema alcanzó elevados niveles de precisión, recall y exactitud en las tareas de detección e identificación, confirmando su capacidad para operar de manera confiable en tiempo real, incluso bajo las condiciones dinámicas y variables propias de los entornos de video. La detección inicial de rostros mediante MTCNN y la generación de embeddings faciales utilizando el modelo InceptionResNetV1 constituyeron componentes clave en el desempeño del sistema.

Asimismo, la integración de los modelos entrenados en una aplicación web, apoyada en cámaras IP y en la arquitectura YOLOv8 para el seguimiento de personas, evidencia el potencial de los enfoques de deep learning para fortalecer los sistemas de vigilancia y control en espacios públicos. No obstante, se identificaron limitaciones asociadas a recursos computacionales, escenarios de iluminación deficiente y a la presencia de obstrucciones faciales parciales, lo que resalta la necesidad de futuras líneas de investigación orientadas a combinaciones de nuevos algoritmos de detección y reconocimiento facial, así como a la robustez del sistema en condiciones adversas.

En conclusión, los resultados obtenidos respaldan la viabilidad de incorporar tecnologías de deep learning en aplicaciones de seguridad pública, abriendo nuevas perspectivas para el desarrollo de sistemas de identificación automática más precisos, eficientes y adaptables a entornos reales.

4.2 Recomendaciones

- Se recomienda realizar pruebas adicionales en entornos con condiciones de iluminación deficiente y variabilidad, para ajustar el rendimiento del sistema en situaciones con baja visibilidad. Estas pruebas permitirán optimizar el modelo en escenarios reales y mejorar su robustez ante obstrucciones faciales parciales y otros desafíos visuales.
- Es fundamental explorar la integración del sistema con cámaras y tecnologías de videovigilancia existentes, asegurando que el modelo YOLOv8 sea escalable y compatible con infraestructuras ya implementadas en terminales grandes. Para ello, se podría considerar el uso de soluciones basadas en la nube o Edge computing para mejorar la eficiencia y la velocidad de procesamiento en tiempo real.
- Se sugiere desarrollar y probar nuevos algoritmos de detección y reconocimiento facial que sean capaces de manejar mejor la variabilidad de las condiciones de operación, como cambios en el entorno, iluminación o personas parcialmente cubiertas. El uso de técnicas híbridas y combinadas podría potenciar aún más la capacidad del sistema.
- Se recomienda investigar la posibilidad de extender el sistema para realizar seguimiento de personas en escenarios de alta afluencia, lo que permitiría mejorar la gestión de seguridad y optimizar la monitorización en tiempo real de sujetos de interés.

- Finalmente, sería beneficioso evaluar el desempeño del sistema en otros contextos de seguridad pública, como en aeropuertos, estaciones de tren y otros espacios con alta densidad de personas. Esto ayudaría a validar la efectividad del sistema en escenarios complejos y contribuiría al desarrollo de sistemas de identificación automática más robustos y adaptativos.

REFERENCIAS

[1] M. H. Khan, M. S. Farid, y M. Grzegorzek, «Vision-based approaches towards person identification using gait», noviembre de 2021, *Elsevier Ireland Ltd*. doi: 10.1016/j.cosrev.2021.100432.

[2] S. Kalfoglou y E. H. Yükseloğlu, «Crime Knows No Borders: The Necessity of International Police Cooperation», *Lectio Soc.*, ago. 2024, doi: 10.47478/lectio.1492496.

[3] Instituto Nacional de Estadística e Informática (INEI), «Informe Técnico - Estadísticas de Seguridad Ciudadana». Accedido: 4 de abril de 2025. [En línea]. Disponible en: https://www.inei.gob.pe/media/MenuRecursivo/boletines/boletin-de-seguridad-ciudadana.pdf

[4] Agencia Andina, «Cusco: Policía elabora banners de “Los Más Buscados” para ayudar a captura | Noticias | Agencia Peruana de Noticias Andina». Accedido: 4 de abril de 2025. [En línea]. Disponible en: https://andina.pe/agencia/Policía Nacional del Perúandina.pe/agencia/noticia-cusco-policia-elabora-banners-los-mas-buscados-para-ayudar-a-captura-699188.aspx

[5] Defensoria del Pueblo, «Chiclayo: Defensoría del Pueblo pide a alcaldes de La Victoria y José Leonardo Ortiz poner énfasis en lucha contra la inseguridad ciudadana - Noticias - Defensoría del Pueblo - Plataforma del Estado Peruano». Accedido: 4 de abril de 2025. [En línea]. Disponible en: https://www.gob.pe/institucion/defensoria/noticias/858575-chiclayo-defensoria-del-pueblo-pide-a-alcaldes-de-la-victoria-y-jose-leonardo-ortiz-poner-enfasis-en-lucha-contra-la-inseguridad-ciudadana

[6] INTERPOL, «Análisis de información policial». Accedido: 4 de abril de 2025. [En línea]. Disponible en: https://www.interpol.int/es/Como-trabajamos/Analisis-de-informacion-criminal/Analisis-de-informacion-policial

[7] F. A. N. M. J. H. V. ARANGOITIA, «EL EXPERTO BIOMÉTRICO DUAL, PARA

ENFRENTAR LA INEFICACIA EN LA COBERTURA DE RESOLUCIÓN DE CASOS CRIMINALES DE IDENTIFICACIÓN POLICIAL, DE HUELLAS DACTILARES E IMÁGENES FACIALES, EN LA DIVISIÓN DE IDENTIFICACIÓN DE LA DIRECCIÓN DE CRIMINALÍSTICA DE LA PNP"», 2020.

[8] la republica, «Urgen más recursos para el sistema de seguridad de la MPCh | Inseguridad ciudadana |Delincuencia | PNP | Sociedad | La República». Accedido: 4 de abril de 2025. [En línea]. Disponible en: https://larepublica.pe/sociedad/2023/08/05/chiclayo-urgen-mas-recursos-para-el-sistema-de-seguridad-de-la-mpch-inseguridad-ciudadana-delincuencia-pnp-418705

[9] D. Chahyati, M. I. Fanany, y A. M. Arymurthy, «Tracking People by Detection Using CNN Features», en *Procedia Computer Science*, Elsevier B.V., 2017, pp. 167-172. doi: 10.1016/j.procs.2017.12.143.

[10] D. A. Maharani, C. MacHbub, P. H. Rusmin, y L. Yulianti, «Real-Time Human Tracking Using Multi-Features Visual With CNN-LSTM and Q-Learning», *IEEE Access*, vol. 12, pp. 13233-13247, 2024, doi: 10.1109/ACCESS.2024.3355785.

[11] T. Matsukawa y E. Suzuki, «Convolutional feature transfer via camera-specific discriminative pooling for person re-identification», en *Proceedings - International Conference on Pattern Recognition*, Institute of Electrical and Electronics Engineers Inc., 2020, pp. 8408-8415. doi: 10.1109/ICPR48806.2021.9412420.

[12] B. Chen, Y. Zha, Y. Li, S. Zhang, Y. Zhang, y T. Ku, «Person Re-Identification Based on Convolutional Neural Network Discriminative Feature Learning», *Guangxue Xuebao/Acta Opt. Sin.*, vol. 38, jul. 2018, doi: 10.3788/AOS201838.0720001.

[13] M. Li, L. Yuan, X. Wen, J. Wang, G. Xie, y Y. Jia, «Multi-Scale Attention Network Based on Multi-Feature Fusion for Person Re-Identification», en *Proceedings of the International Joint Conference on Neural Networks*, Institute of Electrical and Electronics Engineers Inc., jul. 2021. doi: 10.1109/IJCNN52387.2021.9534272.

[14] J. Dhanalakshmi, A. Kumar M, S. J, y S. Devi M, «Identification of Missing Person using CNN», *Int. J. Innov. Sci. Res. Technol.*, pp. 3381-3385, jul. 2024, doi:

10.38124/ijisrt/IJISRT24APR2690.

[15] H. Chen, S. A. Ihnatsyeva, R. P. Bohush, y S. V. Ablameyko, «Person Re-identification in Video Surveillance Systems Using Deep Learning: Analysis of the Existing Methods», *Autom. Remote Control*, vol. 84, n.º 5, pp. 497-528, may 2023, doi: 10.1134/S0005117923050041.

[16] H. Chen, S. Ihnatsyeva, R. Bohush, y S. Ablameyko, «Choice of Activation Function in Convolutional Neural Networks for Person Re-Identification in Video Surveillance Systems», *Program. Comput. Softw.*, vol. 48, n.º 5, pp. 312-321, oct. 2022, doi: 10.1134/S0361768822050036.

[17] J. Chen, Q. Zhang, W.-S. Zheng, y X. Xie, «Efficient and Switchable CNN for Crowd Counting Based on Embedded Terminal», *IEEE Access*, vol. 7, pp. 51533-51541, 2019, doi: 10.1109/ACCESS.2019.2910458.

[18] M. H. Benavides Canchari, M. C. Salcedo Gonzalez, E. Y. Caceres Aldana, W. A. Lazo Tapia, y D. P. Chipana Gago, «Approach to the Luxand Face Facial Recognition System Aimed at the Detection of People in the Criminalistics Unit of the PNP in Huancayo City, Peru», en *Proceedings of the 2023 6th International Conference on Information Science and Systems*, New York, NY, USA: ACM, ago. 2023, pp. 215-223. doi: 10.1145/3625156.3625188.

[19] Piyush Madan y Samaya Madhavan, «Introduction to deep learning», https://developer.ibm.com/learningpaths/get-started-with-deep-learning/an-introduction-to-deep-learning/.

[20] Jesús Bobadilla Sancho, «Machine Learning y Deep Learning». Accedido: 4 de abril de 2025. [En línea]. Disponible en: https://books.google.com.pe/books?id=2c-4EAAAQBAJ&printsec=frontcover&hl=es#v=onepage&q&f=false

[21] J. Bobadilla, *Machine Learning y Deep Learning Usando Python, Scikit y Keras Informática*, Ra-ma (Esp., vol. 1. Bogotá, 2020. Accedido: 4 de abril de 2025. [En línea]. Disponible en: https://api.pageplace.de/preview/DT0400.9789587921465_A41974869/preview-

9789587921465_A41974869.pdf

[22] C. C. Aggarwal, «Educational and software resources for data classification», en *Data Classification: Algorithms and Applications*, CRC Press, 2014, pp. 657-665. doi: 10.1201/b17320.

[23] H. Sánchez Carlessi, C. Reyes Romero, y K. Mejía Sáenz, «Manual de términos en investigación científica, tecnológica y humanística», https://repositorio.urp.edu.pe/entities/publication/fe7f4b74-3863-47e4-b3dc-bd7d52c33bee.

[24] O. Enrique y L. Guevara, «Planteamiento de un algoritmo para la identificación de posibles delincuentes mediante reconocimiento facial en la ciudad de Bogotá», 2020.

[25] Y. C. Quispe Quispe, *Diseños y Secuencia Didáctica para la Investigación en un Nuevo Paradigma*. CID - Centro de Investigación y Desarrollo, 2023. doi: 10.37811/cli_w957.

[26] K. Zhang, Z. Zhang, Z. Li, y Y. Qiao, «Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks», *IEEE Signal Process. Lett.*, vol. 23, pp. 1499-1503, oct. 2016, doi: 10.1109/LSP.2016.2603342.

[27] P. Schäfer, J. Brand, U. Leser, B. Peng, y T. Palpanas, «Fast and Exact Similarity Search in less than a Blink of an Eye», nov. 2024.

[28] P. Jiang, D. Ergu, F. Liu, Y. Cai, y B. Ma, «A Review of Yolo Algorithm Developments», en *Procedia Computer Science*, Elsevier B.V., 2021, pp. 1066-1073. doi: 10.1016/j.procs.2022.01.135.

[29] «Gente cruzando la calle en un: video de stock (totalmente libre de regalías) 1097693717 | Shutterstock», https://www.shutterstock.com/es/video/clip-1097693717-people-crossing-street-crosswalk-green-traffic-light.

[30] P. D. N. H. Sai, V. V. S. Kiran, K. Rohith, y D. Rajeswararao, «Identification of Missing Person Using Convolutional Neural Networks», *Int. Conf. Sustain. Comput. Data Commun. Syst. ICSCDS 2022 - Proc.*, pp. 485-489, 2022, doi: 10.1109/ICSCDS53736.2022.9760777.

ANEXOS

Anexo 1. Acta de revisión de similitud de la investigación.

Universidad Señor de Sipán

ACTA DE REVISIÓN DE SIMILITUD DE LA INVESTIGACIÓN

Yo del curso de **Investigación II**, del Programa de Estudios de **Ingeniería de Sistemas**, luego de revisar la investigación de los estudiantes, Mondragon Fernandez Ale, Yarango Farro Darwin Orlando, titulada:

Redes neuronales convolucionales para la identificación automática de individuos sujetos a búsqueda en terminales terrestres

Dejo constancia que la investigación antes indicada tiene un índice de similitud del porcentaje 11%, verificable en el reporte de originalidad mediante el software de similitud TURNITIN. Por lo que se concluye que cada una de las coincidencias detectadas no constituyen plagio y cumple con lo establecido en la Directiva sobre índice de similitud de los productos académicos y de investigación en la Universidad Señor de Sipán S.A.C. vigente.

En virtud de lo antes mencionado, firma:

William Joel Marín Rodriguez	DNI: 40398872	

Pimentel, 13 de diciembre del 2024

Anexo 2. Acta de aprobación de asesor.

Universidad
Señor de Sipán

ACTA DE APROBACIÓN DEL ASESOR

Yo **Heber Ivan Mejia Cabrera**, quien suscribe como asesor designado mediante Resolución de Facultad N° **0534-2024/FIAU-USS**, del proyecto de investigación titulado Redes neuronales convolucionales para la identificación automática de individuos sujetos a búsqueda en terminales terrestres, desarrollado por los estudiantes: **Mondragón Fernández Alex , Yarango Farro Darwin Orlando**, del programa de estudios de **Ingeniería de Sistemas**, acredito haber revisado, y declaro expedito para que continue con el trámite pertinentes.

En virtud de lo antes mencionado, firma:

Heber Ivan Mejia Cabrera	DNI: 41639565	

Pimentel, 16 de diciembre de 2024

Anexo 3. Correo de recepción del manuscrito remitido por la revista.

IEEE Access <onbehalfof@manuscriptcentral.com> 11 abr 2025, 2:

para mí, mfernandezalex

inglés → español
Mostrar original

11 de abril de 2025

Acceso-2024-50486 - Redes neuronales convolucionales para la identificación automática de individuos en terminales terrestres

Estimado señor Farro:

Felicitaciones por su artículo recientemente aceptado en IEEE Access.

Continúe enviando sus archivos finales a través del Portal de autores del IEEE .

Si tiene alguna pregunta sobre los materiales requeridos o el proceso de envío final, no dude en ponerse en contacto conmigo.

Atentamente,

Señor Nishant Shukla
IEEE Access

Responder | Responder a todos | Reenviar

Anexo 4. Tabla de operacionalización de variables.

Variable	Definición Conceptual	Definición Operacional	Dimensiones	Indicadores	Ítems	Instrumento	Valores Finales	Tipo de Variable	Escala de Medición
Identificación de individuos sujetos a búsqueda en terminales terrestres.	Grado en que el sistema de identificación es capaz de reconocer correctamente a los individuos buscados en terminales terrestres.	Evaluación del rendimiento del sistema de identificación basado en CNN mediante la medición de la tasa de verdaderos positivos y la tasa de falsos positivos en un conjunto de datos de prueba.	Identificación	. Tasa de verdaderos positivos (TP) . Tasa de falsos positivos (FP)	. TP = N° de aciertos en identificación . FP = N° de errores por confusión	- Ficha de Observación. - Matriz de confusión.	- % de TP - % de FP	Dependiente	Escala de razón

Variable	Definición Conceptual	Definición Operacional	Dimensiones	Indicadores	Ítems	Instrumento	Valores Finales	Tipo de Variable	Escala de Medición
Redes Neuronales Convolucionales (CNN).	Modelo de aprendizaje profundo utilizado para el reconocimiento de patrones faciales en imágenes.	Implementación de modelos de red neuronal convolucional (CNN) entrenado para la identificación de individuos en imágenes capturadas mediante técnicas de machine learning, evaluado a través de métricas.	Tecnología de machine learning.	- Precisión. - Exactitud. - Recall. - F1 Score	.P=TP/(TP+FP) .ACC=(TP+TN)/(TP+TN+FP+ FN) .R=TP/(TP+FN) .F=2*(P*R)/(P+ R)	- Matriz de confusión.	- 0 – 100%.	Independiente	Porcentaje

Anexo 5. Código fuente del sistema de identificación.

https://github.com/Alexmondra/FaceGuard

Anexo 6. Aplicación web

El proceso inicia con el registro en la plataforma web, los usuarios autorizados, como personal de seguridad y operadores de vigilancia, pueden acceder al sistema mediante credenciales únicas (Figura 10).

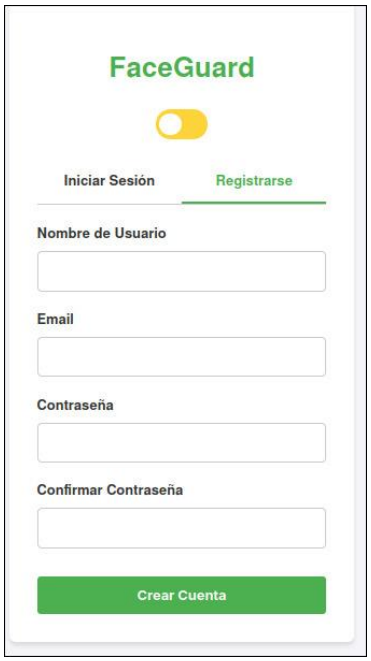

Figura 10. Registro al sistema

Fuente: Elaboración del autor

Una vez registrado, el usuario accede al sistema proporcionando sus credenciales (correo electrónico y contraseña) en el formulario de inicio de sesión. El sistema valida la información ingresada y, si es correcta, redirige al Panel Principal (Figura 11).

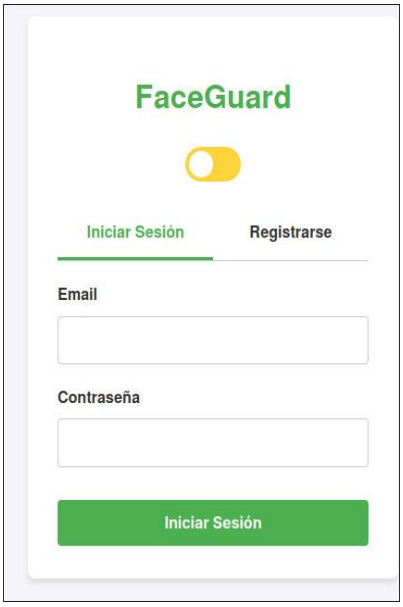

Figura 11. Inicio de sesión al sistema

Fuente: Elaboración del autor

En la Figura 12, se muestra el módulo de registro de nuevas personas sujetas a búsqueda. Este apartado permite a los usuarios autorizados ingresar los datos relevantes de las personas a identificar, incluyendo nombre, características físicas, fotografía y cualquier información adicional que facilite su localización.

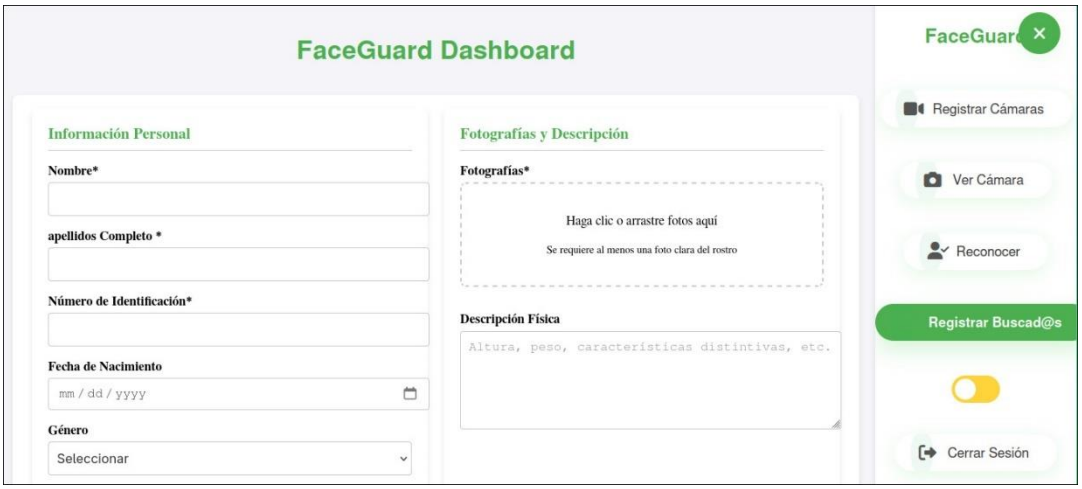

Figura 12. Registro de personas

Fuente: Elaboración del autor

En la Figura 13, se visualiza el listado de todas las personas registradas para ser identificadas, mostrando las imágenes y la información previamente ingresada durante el registro.

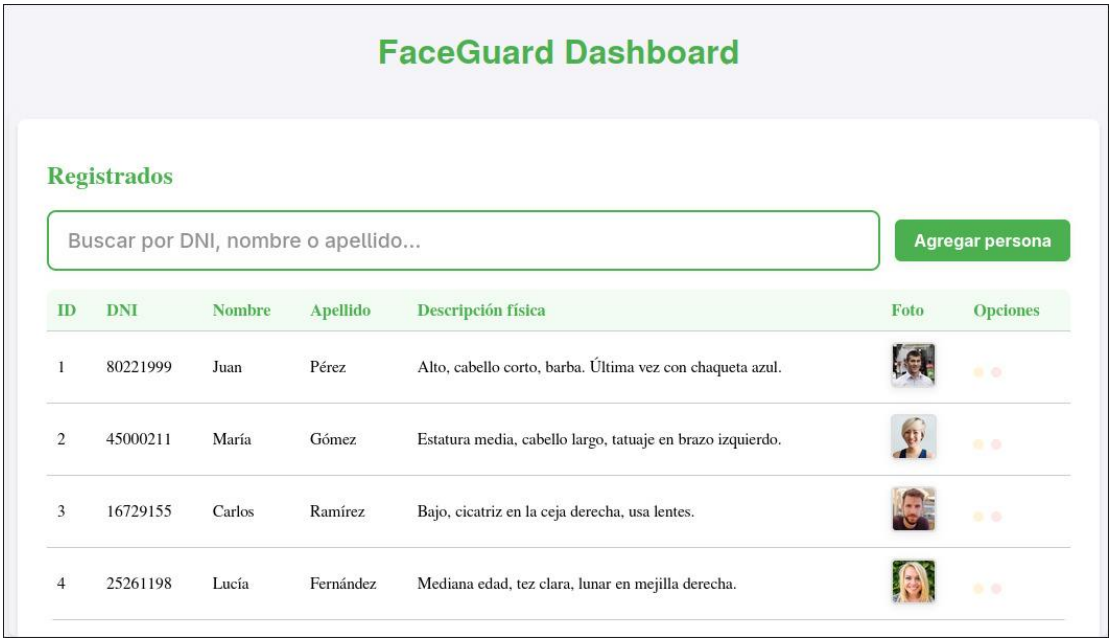

Figura 13. Lista de personas a identificar

Fuente: Elaboración del autor

El usuario puede subir una imagen para verificar si la persona está registrada como buscada. El sistema compara automáticamente la imagen con los registros existentes y muestra si la persona ha sido registrada previamente (Figura 14).

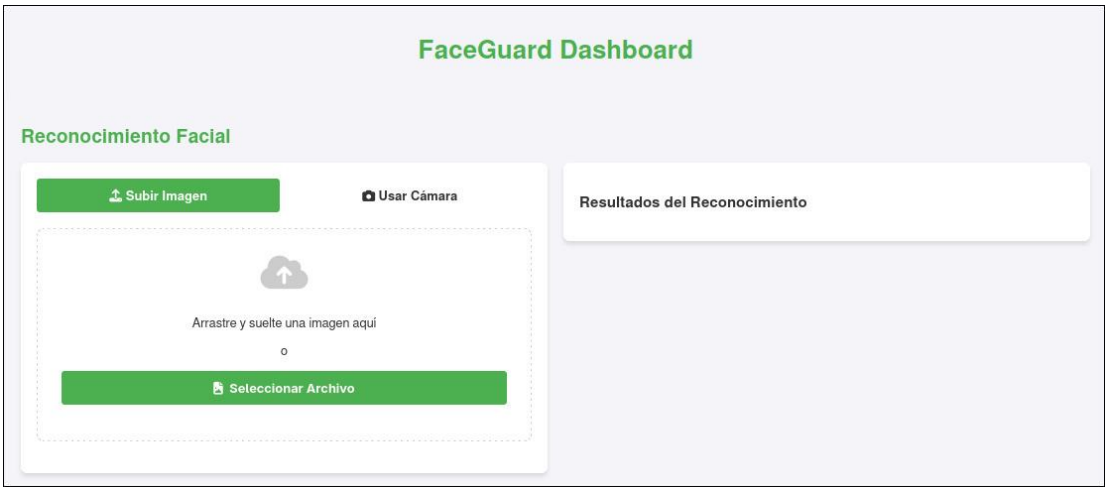

Figura 14. Reconocimiento de personas

Fuente: Elaboración del autor

En la figura 15 se muestra el proceso de registro de nuevas cámaras permite agregar dispositivos de vigilancia al sistema. Cada cámara se registra con un ID único del usuario que garantiza que solo ese usuario autorizado tenga acceso a las imágenes y grabaciones capturadas por dichas cámaras.

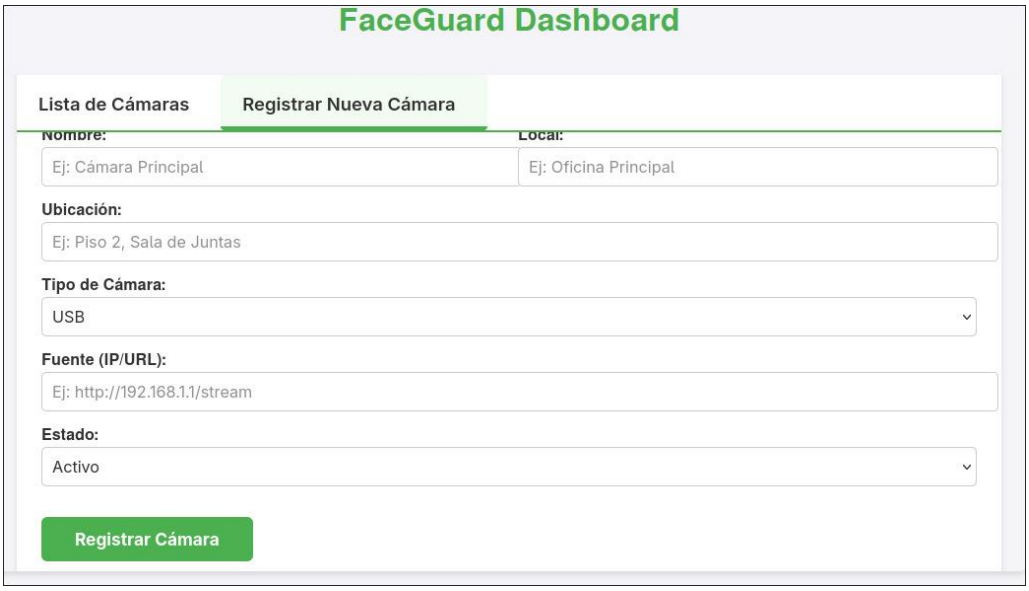

Figura 15. Registro de cámaras

Fuente: Elaboración del autor

El usuario puede seleccionar las cámaras registradas que desea visualizar. Según su selección, el sistema le permitirá acceder a las transmisiones en vivo de las cámaras elegidas (Figura 16).

Figura 16. Selección de cámaras

Fuente: Elaboración del autor

Cuando el usuario selecciona las cámaras a visualizar, el sistema le permite ver las transmisiones en vivo de las cámaras elegidas, asegurando una visualización en tiempo real de las áreas de interés (Figura 17).

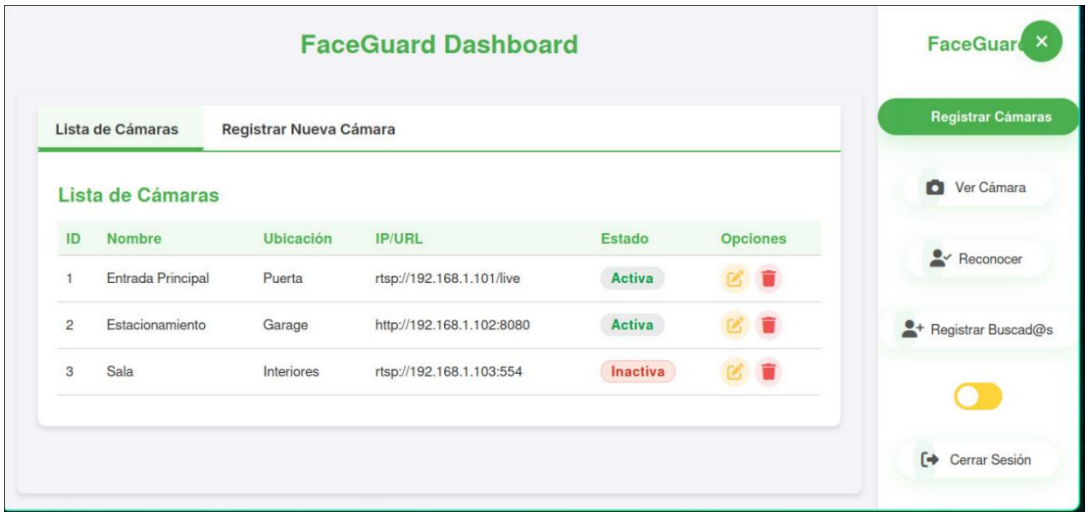

Figura 17. Listad de cámaras registra por el usuario

Fuente: Elaboración del autor