UNIVERSIDAD POLITÉCNICA ESTATAL DEL CARCHI

FACULTAD DE INDUSTRIAS AGROPECUARIAS Y CIENCIAS AMBIENTALES

CARRERA DE COMPUTACIÓN

Tema: "Reconocimiento facial en circuito cerrado de video vigilancia."

Trabajo de Integración Curricular previo a la obtención del título de Ingeniero en Ciencias de la Computación

AUTOR: Ayala Acosta Erik Gustavo

TUTOR: Ing. Milton Gabriel Del Hierro Mosquera MSc

Tulcán, 2025.

CERTIFICADO DEL TUTOR

Certifico que el estudiante Ayala Acosta Erik Gustavo con el número de cédula 0402117881 respectivamente ha desarrollado el Trabajo de Integración Curricular: "Reconocimiento facial en circuito cerrado de video vigilancia."

Este trabajo se sujeta a las normas y metodología dispuesta en la Codificación del Reglamento de Régimen Académico y de Estudiantes de la UPEC, por lo tanto, autorizo la presentación de la sustentación para la calificación respectiva.

Firmado electrónicamente por:
MILTON GABRIEL DEL HIERRO MOSQUERA
Validar únicamente con FirmaEC

_____

Ing. Milton Gabriel Del Hierro Mosquera MSc.

TUTOR

Tulcán, noviembre de 2025

AUTORÍA DE TRABAJO

El presente Trabajo de Integración Curricular constituye un requisito previo para la obtención del título de Ingeniero en la Carrera de Computación de la Facultad de Industrias Agropecuarias y Ciencias Ambientales

Yo, Ayala Acosta Erik Gustavo con cédula de identidad número 0402117881 y respectivamente declaro que la investigación es absolutamente original, auténtica, personal y los resultados y conclusiones a los que he llegado son de mi absoluta responsabilidad.

Ayala Acosta Erik Gustavo
AUTOR

Tulcán, noviembre de 2025

ACTA DE CESIÓN DE DERECHOS DEL TRABAJO DE INTEGRACIÓN CURRICULAR

Yo Ayala Acosta Erik Gustavo declaro ser autor de los criterios emitidos en el Trabajo de Integración Curricular: "Reconocimiento facial en circuito cerrado de video vigilancia." y eximo expresamente a la Universidad Politécnica Estatal del Carchi y a sus representantes de posibles reclamos o acciones legales.

Ayala Acosta Erik Gustavo

AUTOR

Tulcán, noviembre de 2025

AGRADECIMIENTO

Quiero expresar mi más sincero agradecimiento a todas las personas que hicieron posible la culminación de este trabajo de titulación, el cual representa no solo un logro académico, sino también personal y profesional.

En primer lugar, agradezco profundamente a Dios, por brindarme la fortaleza, sabiduría y perseverancia necesarias para superar cada desafío a lo largo de este proceso.

A mis padres y familia, por su amor incondicional, su apoyo constante y sus palabras de aliento, que fueron el pilar fundamental para mantenerme enfocado en alcanzar esta meta.

A mi tutor de tesis, por su acompañamiento, orientación técnica y por compartir su conocimiento durante el desarrollo de esta investigación. Su guía fue clave para consolidar este proyecto basado en el reconocimiento facial aplicado a la seguridad universitaria.

De igual manera, agradezco a mis docentes y compañeros de la carrera de Computación, por su apoyo, cooperación y consejos que enriquecieron mi formación académica.

Finalmente, extiendo mi gratitud a la Universidad Politécnica Estatal del Carchi, por brindarme las herramientas, los conocimientos y el entorno académico necesario para desarrollar este proyecto.

DEDICATORIA

A Dios,
por ser mi guía y darme la fortaleza para avanzar con fe en cada etapa de mi vida académica.

A mis padres,
por su amor, comprensión y sacrificio, que me inspiraron a esforzarme y alcanzar mis metas.

A mi familia,
por su apoyo incondicional y su confianza constante en mi capacidad para lograr este objetivo.

A mis docentes,
por compartir sus conocimientos y motivarme a continuar aprendiendo con pasión y compromiso.

A la Universidad Politécnica Estatal del Carchi,
por ser el espacio donde crecí personal y profesionalmente, y que me permitió desarrollar este proyecto.

ÍNDICE

ÍNDICE DE TABLAS

ÍNDICE DE FIGURAS

ÍNDICE DE ANEXOS

RESUMEN

El presente proyecto de investigación desarrolló un sistema de reconocimiento facial integrado al circuito cerrado de videovigilancia de la Carrera de Computación de la Universidad Politécnica Estatal del Carchi, con el objetivo de mejorar la seguridad mediante identificación automatizada en tiempo real. Se analizaron y seleccionaron tecnologías de reconocimiento facial, determinando que FaceNet representa la solución más adecuada por su capacidad de generar embeddings robustos de 128 dimensiones mediante redes neuronales convolucionales. Se aplicaron exitosamente algoritmos de reconocimiento facial sobre el servidor HPE ProLiant DL360 Gen10 existente, demostrando capacidad de escalabilidad de usuarios. Se propuso e implementó un sistema completo basado en el modelo FaceNet integrado con cámaras IP Hikvision, operando con Ubuntu 22.04.3 LTS y procesando video mediante protocolo RTSP. Se utilizó la metodología Top-Down basada en redes, que permitió una planificación estructurada desde el diseño lógico hasta la configuración física del sistema. La implementación empleó tecnologías de software libre como OpenCV, TensorFlow y Dlib para procesamiento de imágenes y extracción de características faciales. Las pruebas realizadas con cuatro usuarios registrados alcanzaron un accuracy del 63.6% con umbral de confianza de 0.6 en similitud coseno, identificando como principal limitación la ausencia de GPU dedicada para aceleración del procesamiento. Los resultados de las encuestas aplicadas a 177 estudiantes mostraron alta aceptación del sistema (77% a favor), destacando mejoras percibidas en precisión del control de acceso y confianza en la seguridad. Se proponen tres alternativas de mejora: implementación de GPU NVIDIA RTX A2000 (inversión $750, mejora proyectada 85-88% accuracy), actualización de infraestructura completa ($2,100, accuracy 90-92%), y actualización de cámaras a 8MP con IA integrada más GPU ($3,900, accuracy proyectado 93-95%). El proyecto demuestra la viabilidad técnica y aceptación social del reconocimiento facial en entornos académicos.

Palabras Claves: Reconocimiento facial, Videovigilancia, FaceNet, OpenCV, TensorFlow, Embeddings faciales, Deep Learning.

ABSTRACT

This research project developed a facial recognition system integrated into the closed-circuit video surveillance network of the Computer Science Program at the Universidad Politécnica Estatal del Carchi, with the goal of enhancing security through real-time automated identification. Facial recognition technologies were analyzed and evaluated, determining that FaceNet represented the most suitable solution due to its ability to generate robust 128 dimensional embeddings using convolutional neural networks. Facial recognition algorithms were successfully applied on the existing HPE ProLiant DL360 Gen10 server, demonstrating user scalability. A complete system based on the FaceNet model was proposed and implemented, integrated with Hikvision IP cameras, running on Ubuntu 22.04.3 LTS, and processing video streams through the RTSP protocol. The network-based Top-Down methodology was used to ensure a structured planning process from logical design to physical configuration. The implementation employed open-source technologies such as OpenCV, TensorFlow, and Dlib for image processing and facial feature extraction. Tests conducted with four registered users achieved an accuracy of 63.6% with a confidence threshold of 0.6 in cosine similarity, identifying the main limitation as the lack of a dedicated GPU for processing acceleration. Surveys administered to 177 students indicated high acceptance of the system (77% favorable), highlighting perceived improvements in access control accuracy and overall security confidence. Three improvement alternatives were proposed: the implementation of an NVIDIA RTX A2000 GPU (investment $750, projected 85–88% accuracy), full infrastructure upgrade ($2,100, 90–92% accuracy), and camera upgrades to 8MP with integrated AI plus GPU support ($3,900, projected 93–95% accuracy). The project demonstrates both the technical feasibility and social acceptance of facial recognition systems in academic environments.

Keywords: Facial recognition, Video surveillance, FaceNet, OpenCV, TensorFlow, Facial embeddings, Deep Learning.

INTRODUCCIÓN

En el contexto universitario actual, garantizar la seguridad física y patrimonial de las instalaciones se ha convertido en una prioridad, especialmente en áreas que albergan equipos tecnológicos de alto valor y datos sensibles, como los laboratorios de computación. Las instituciones de educación superior son entornos dinámicos con un flujo constante de estudiantes, docentes, personal administrativo y visitantes, lo que genera una necesidad crítica de implementar mecanismos de control de acceso eficientes y confiables (Llerena Yupanqui & La Madrid Aliaga, 2022). Sin embargo, muchos de los sistemas de vigilancia tradicionales, basados en métodos manuales como guardias de seguridad o registros en papel, presentan limitaciones significativas, como la propensión a errores humanos, la posibilidad de suplantación de identidad y la incapacidad para responder de manera inmediata ante incidentes de seguridad (Ayala Heredia, 2021).

Frente a estas problemáticas, la tecnología de reconocimiento facial emerge como una solución innovadora dentro de los sistemas de circuito cerrado de televisión (CCTV), permitiendo la identificación automatizada y precisa de individuos en tiempo real. Esta tecnología no solo optimiza los procesos de control de acceso, sino que también fortalece la percepción de seguridad dentro de la comunidad universitaria (Bajaña Ortiz, 2023). Su aplicación ha sido explorada en diversos contextos educativos, demostrando potencial para reducir accesos no autorizados y agilizar la gestión de la seguridad (Galindo Taype et al., 2021).

En el caso específico de la Carrera de Computación de la Universidad Politécnica Estatal del Carchi (UPEC), se ha identificado la necesidad de modernizar su sistema de videovigilancia. Actualmente, el acceso a los laboratorios y centros tecnológicos carece de un mecanismo automatizado de identificación, lo que representa un riesgo latente para los recursos institucionales y la integridad de quienes hacen uso de estos espacios (Ayala Heredia, 2021). Por ello, este proyecto de investigación se centra en el diseño e implementación de un sistema de reconocimiento facial integrado al circuito

de videovigilancia existente, utilizando algoritmos de aprendizaje profundo como FaceNet y herramientas de software libre como OpenCV y TensorFlow.

El sistema propuesto busca realizar una verificación en tiempo real de las personas que acceden a las áreas controladas, cotejando sus rostros con una base de datos previamente registrada. De esta forma, se espera no solo disuadir y prevenir ingresos no autorizados, sino también generar un entorno más seguro y controlado, acorde con los avances tecnológicos contemporáneos (Navarro-Dolmestch, 2023). Además, se han considerado aspectos éticos y legales relacionados con el manejo de datos biométricos, asegurando que la implementación se realice dentro del marco normativo aplicable.

La presente investigación se estructura en siete capítulos. El primero aborda el planteamiento y formulación del problema, así como la justificación y los objetivos. El segundo capítulo desarrolla el marco teórico, revisando antecedentes y fundamentos conceptuales sobre inteligencia artificial, visión por computadora y reconocimiento facial. El tercer capítulo detalla la metodología empleada, con un enfoque mixto y un diseño basado en la metodología Top-Down. El cuarto capítulo presenta los resultados obtenidos y su discusión, mientras que el quinto expone las conclusiones y recomendaciones derivadas del estudio. Finalmente, el sexto y séptimo capítulo contienen las referencias bibliográficas y los anexos del proyecto.

I. EL PROBLEMA

1.1. PLANTEAMIENTO DEL PROBLEMA

En la actualidad, la seguridad en las instituciones de educación superior se ha convertido en un aspecto fundamental, debido a la necesidad de garantizar que tanto los estudiantes como el personal se desarrollen en un entorno seguro y controlado. En diversos países se han implementado tecnologías avanzadas, entre ellas los sistemas de videovigilancia basados en inteligencia artificial, especialmente aquellos que utilizan reconocimiento facial. Estas herramientas han demostrado ser altamente efectivas para la identificación rápida de personas y la prevención de incidentes dentro de los campus universitarios. De hecho, en países como Estados Unidos, China y el Reino Unido, su aplicación en el ámbito académico ha permitido fortalecer los mecanismos de control y mejorar el análisis de seguridad en tiempo real (Ragas, 2020).

En el contexto ecuatoriano, aún se evidencia la necesidad de mejorar los sistemas de seguridad universitarios. El avance tecnológico actual exige soluciones más modernas que posibiliten el procesamiento ágil de grandes volúmenes de información. Sin embargo, como señalan Albán Gómez y Pilay Ríos (2024), muchas universidades del país continúan utilizando métodos tradicionales que resultan insuficientes para las exigencias contemporáneas, especialmente cuando se requiere monitoreo y control en tiempo real.

La Universidad Politécnica Estatal del Carchi no constituye una excepción a esta realidad. En las carreras de Computación, los procedimientos de vigilancia aún dependen en gran medida de la intervención humana, como la observación manual de las cámaras o el registro en formato físico. Estas prácticas incrementan el margen de error, demandan mayor tiempo y dificultan la respuesta ante posibles eventualidades. La falta de automatización genera vulnerabilidades que pueden ser aprovechadas por personas no autorizadas, poniendo en riesgo tanto la integridad de los individuos como los recursos tecnológicos disponibles en las aulas y laboratorios.

Uno de los principales desafíos radica en la identificación constante y precisa de las personas que ingresan a las instalaciones. Los sistemas convencionales presentan limitaciones de desempeño ante variaciones lumínicas, cambios de ángulo o alta densidad de movimiento, lo que conlleva a pérdidas de información relevante y a un control deficiente. Adicionalmente, las cámaras generan grandes volúmenes de datos que, sin un sistema automatizado capaz de procesarlos en tiempo real, resultan imposibles de analizar oportunamente. La ausencia de un mecanismo eficiente de análisis impide una respuesta inmediata ante eventos de riesgo, lo que podría comprometer la seguridad institucional y la reputación de la universidad.

Ante esta problemática, el presente proyecto plantea la implementación de un sistema de videovigilancia con reconocimiento facial en los laboratorios y espacios tecnológicos de la carrera de Computación. Su propósito es optimizar el control de acceso, automatizar los procesos de monitoreo y reducir la dependencia de la intervención humana. No obstante, la implementación de esta tecnología implica considerar ciertos factores críticos, como la inversión en equipamiento especializado, las limitaciones de infraestructura tecnológica y, de manera prioritaria, la gestión responsable de la privacidad y protección de los datos personales. Con una adecuada planificación técnica y organizacional, se espera que estos desafíos puedan ser abordados eficazmente, garantizando una solución sostenible y segura.

1.2. FORMULACIÓN DEL PROBLEMA

La necesidad de fortalecer el sistema de seguridad existente mediante el uso de reconocimiento facial en el circuito cerrado de videovigilancia de la Carrera de Computación de la Universidad Politécnica Estatal del Carchi.

1.3. JUSTIFICACIÓN

La implementación de un sistema de reconocimiento facial integrado al circuito cerrado de videovigilancia de la Universidad Politécnica Estatal del Carchi (UPEC) se fundamenta en la necesidad imperativa de fortalecer los mecanismos de seguridad en los centros tecnológicos de la Carrera de Computación. Actualmente, la ausencia de restricciones en el ingreso permite el acceso indiscriminado de personal no autorizado, incrementando significativamente los riesgos de sustracción de equipamiento, uso inadecuado de instalaciones y perturbación de actividades académicas programadas.

Los sistemas de videovigilancia convencionales presentan limitaciones operativas inherentes a su dependencia de vigilancia humana continua, la cual es susceptible a factores como fatiga cognitiva, pérdida de concentración y errores de percepción. En contraste, el reconocimiento facial mediante algoritmos de inteligencia artificial constituye una solución tecnológicamente superior que permite la identificación biométrica automatizada con alta precisión en tiempo real, garantizando que únicamente individuos previamente registrados y autorizados puedan acceder a áreas restringidas. Esta tecnología no solo optimiza significativamente los niveles de seguridad, sino que también agiliza los procesos de acceso mediante autenticación automática, reduciendo tiempos de espera y minimizando interrupciones en el flujo operativo institucional.

El sistema propuesto beneficiará directamente a toda la comunidad académica de la Carrera de Computación, incluyendo estudiantes, docentes, personal administrativo y de servicios. La totalidad de usuarios autorizados será registrada en una base de datos, estableciendo un entorno institucional caracterizado por mayor seguridad perimetral y control de accesos estandarizado. Adicionalmente, la adopción de esta tecnología emergente posicionará estratégicamente a la UPEC como institución líder en innovación tecnológica dentro del contexto universitario ecuatoriano, incrementando su atractivo para estudiantes y profesionales que valoran entornos académicos tecnológicamente avanzados y alineados con las tendencias contemporáneas en seguridad institucional.

1.4. OBJETIVOS Y PREGUNTAS DE INVESTIGACIÓN

1.4.1. Objetivo General

Desarrollar un sistema de reconocimiento facial en el circuito cerrado de video vigilancia que posee la carrera de Computación de la Universidad Politécnica Estatal del Carchi.

1.4.2. Objetivos Específicos

- Analizar tecnologías de reconocimiento facial para el desarrollo de un sistema automatizado de verificación de identidad.
- Aplicar algoritmos de reconocimiento facial en los dispositivos electrónicos disponibles en la carrera de Computación con la finalidad de validar su funcionamiento técnico y compatibilidad.

- Proponer un sistema de reconocimiento facial en el circuito cerrado de video vigilancia con el propósito de implementar un modelo de generación de *vectores de características faciales (embeddings)* que permita identificar de forma automática a los usuarios

1.4.3. Preguntas de Investigación

- ¿Cuáles son los algoritmos de reconocimiento facial más adecuados?
- ¿Cuáles son las preocupaciones de privacidad más relevantes para el personal del laboratorio en relación con la recopilación y el procesamiento de datos biométricos para el reconocimiento facial?
- ¿Cómo afectan las condiciones ambientales, como la iluminación y la distancia a la cámara, a la precisión y la velocidad del reconocimiento facial?

II. FUNDAMENTACIÓN TEÓRICA

2.1. ANTECEDENTES DE LA INVESTIGACIÓN

En los últimos años, el reconocimiento facial ha avanzado mucho y hoy en día se ha vuelto una herramienta clave en temas de seguridad y vigilancia. Su uso en sistemas de videovigilancia (CCTV) ha demostrado ser bastante útil para mejorar la seguridad, incluso en lugares como colegios y universidades. Por eso, en esta parte se revisan algunos estudios recientes que hablan sobre cómo se ha aplicado esta tecnología.

En el caso de la Universidad Politécnica Estatal del Carchi, actualmente el acceso a los laboratorios de las carreras de Computación e Ingeniería en Informática todavía se hace de forma manual, lo cual representa varios problemas en cuanto a seguridad y control. Para solucionar esto, se desarrolló una investigación llamada "*Control y registro de personal mediante el uso de las TIC para el acceso a la Universidad Politécnica Estatal del Carchi en el periodo 2019-2020*". En ese estudio se propuso un sistema de control de acceso usando reconocimiento facial.

El sistema se diseñó con varias herramientas tecnológicas: Python se usó para los algoritmos, OpenCV para detectar los rostros, SQLite como base de datos, y un Raspberry Pi 3 Modelo B con su sistema operativo (Raspberry Pi OS), que se encarga de abrir la puerta solo a las personas que están registradas. Además, el proyecto se hizo usando una metodología mixta, es decir, se combinaron encuestas y revisión teórica con el desarrollo de un prototipo. Esto permitió crear un sistema que se ajusta a las necesidades de la universidad, buscando modernizar el acceso a los centros tecnológicos, hacerlo más eficiente y sobre todo, más seguro(Ayala Heredia, 2021).

Un estudio referencial realizado en Colombia, enfocado en la integración de tecnologías de reconocimiento facial en los sistemas de Circuito Cerrado de Televisión (CCTV), destacó la relevancia de adaptar estas tecnologías a las condiciones culturales y socioeconómicas del país. La investigación, llevada a cabo por Castañeda Rincón y Santos Ariza (2021), analizó experiencias internacionales, especialmente en Europa y América, para identificar las mejores prácticas y aspectos técnicos.

Según la investigación de Llerena Yupanqui & La Madrid Aliaga (2022), hay tres razones principales por las que los sistemas de videovigilancia no funcionan tan bien como deberían. Primero, no se están usando bien los parámetros para decidir dónde colocar las cámaras. Segundo, la calidad de las imágenes que capturan no es la mejor. Y tercero, hay problemas en cómo se gestiona el personal encargado del reconocimiento facial digital.

Para enfrentar estos problemas, los autores proponen una idea bastante útil: crear una guía que sirva para estandarizar tanto las especificaciones técnicas como la ubicación ideal de las cámaras. Con esto se espera mejorar la calidad de las imágenes, lo que también facilitaría el trabajo de los especialistas encargados de analizar los videos, especialmente en casos legales dentro de la ciudad de Lima.

El uso del Reconocimiento Facial Automático (AFR) por parte de la policía ha generado bastante polémica, sobre todo por cómo podría afectar los derechos de las personas, en especial la privacidad y el manejo de los datos personales. Un caso reciente que revisó el Alto Tribunal de Justicia de Inglaterra y Gales mostró que, aunque esta tecnología puede meterse un poco en la vida privada de la gente, en algunos casos se considera válida y justificada, siempre que se sigan ciertas reglas legales.

Eso sí, el estudio también tocó un tema delicado: el posible sesgo en el sistema. Se revisó si el AFR afectaba más a mujeres o a personas de minorías raciales. La conclusión fue que no hay pruebas claras de que eso ocurra de forma significativa. Aun así, esto demuestra que es necesario tener una regulación bien pensada que encuentre el equilibrio entre mantener la seguridad pública y respetar los derechos de las personas. Y este es un tema que también vale la pena considerar dentro de las leyes en Perú (Flores, 2021).

Durante la pandemia del COVID-19, la Universidad Nacional de Educación a Distancia (UNED), en España, tuvo que ajustarse a las restricciones y buscar formas nuevas para seguir evaluando a los estudiantes desde sus casas. Por eso empezaron a usar sistemas de reconocimiento facial durante los exámenes virtuales, con la idea de asegurar que todo fuera justo y que la persona que rendía el examen fuera realmente quien decía ser. Aunque fue una solución innovadora, no estuvo libre de problemas.

Desde el lado técnico, los algoritmos usados para el reconocimiento facial no eran perfectos. A veces cometían errores o mostraban cierto sesgo, lo que podía hacer que no identificaran bien a los estudiantes. Eso podía generar situaciones injustas. Además, en el aspecto legal, surgieron varias dudas sobre el uso de datos personales, especialmente por lo que dice el Reglamento General de Protección de Datos (RGPD) de la Unión Europea, que protege la privacidad de las personas.

También hubo preocupaciones éticas. Algunos estudiantes se sintieron incómodos por la vigilancia constante mientras daban sus exámenes, y también por el hecho de que tal vez no se les explicó del todo bien cómo se usarían sus datos. Eso afectó un poco la relación de confianza entre ellos y la universidad.

Con todo esto en cuenta, la UNED —y otras instituciones educativas también— se enfrentan al reto de analizar con mucho cuidado si vale la pena seguir usando este tipo de tecnología. Es importante encontrar un equilibrio entre garantizar la seguridad en los procesos educativos y respetar los derechos y la privacidad de los estudiantes (Aznarte et al., 2022).

El trabajo de Lorenzana Ramos, (2022) se enfocó en crear un sistema que usa visión por computadora para reconocer tanto rostros como emociones, con el fin de que pudiera interactuar con un robot con cara animada en la Universidad del Valle de Guatemala. Para lograrlo, se usaron varias herramientas: Keras fue clave para el modelo de reconocimiento usando aprendizaje profundo, OpenCV ayudó con la parte visual, y Kivy sirvió para crear la interfaz gráfica con la que los usuarios podían interactuar.

El sistema estaba pensado para reaccionar según la emoción que detectara en la persona, lo que hacía que la interacción entre humano y robot fuera más natural. Para que la detección fuera precisa, se estudiaron expresiones faciales que mostraran diferentes emociones y se eligió un conjunto de datos adecuado que representara bien esos estados de ánimo.

El modelo fue entrenado en Python usando redes neuronales que se iban ajustando paso a paso. Primero se usó un conjunto pequeño de imágenes para ir afinando el entrenamiento, y luego se aplicó uno más grande para mejorar la precisión final del sistema. Para el reconocimiento facial, se empleó un clasificador Haar Cascade, pero con modificaciones, y se necesitaba una cámara conectada para que el sistema pudiera funcionar correctamente.

Por último, se diseñó una interfaz en Kivy donde se podía ver la información capturada y cómo respondía el robot, lo cual hizo que la interacción fuera más sencilla y eficiente para el usuario.

En la Universidad Continental se dieron cuenta de que la suplantación de identidad durante los exámenes finales es un problema que pasa con bastante frecuencia, sobre todo en las materias generales, y eso genera preocupación todos los años. Por eso, Galindo Taype y su equipo (2021) realizaron una investigación con el objetivo de crear un sistema de escritorio que pudiera reconocer los rostros de los estudiantes y confirmar su identidad durante los exámenes presenciales, en la sede de Huancayo.

Para desarrollar el sistema, usaron la metodología Kanban y organizaron el trabajo con la herramienta Trello, lo que les ayudó a tener un mejor control de todas las tareas que se iban haciendo. También aplicaron encuestas a estudiantes y profesores para recoger información útil.

Después, armaron un modelo de reconocimiento facial y lo probaron con cinco estudiantes, tomando 50 fotos de cada uno en diferentes situaciones, como con o sin mascarilla, y con o sin lentes. Con esas imágenes armaron una base de datos para hacer las pruebas.

Al final, al evaluar los resultados con una matriz de confusión, se obtuvo un nivel de precisión del 93% en la similitud del reconocimiento facial. Esto demostró que el sistema funciona bien y puede ser una buena herramienta para evitar que alguien se haga pasar por otro en los exámenes(Galindo Taype et al., 2021).

En la investigación de Bajaña Ortiz, (2023), se busca mejorar la seguridad de los estudiantes en la Universidad Técnica de Babahoyo analizando diferentes algoritmos de reconocimiento facial. El propósito principal es encontrar los lugares más adecuados dentro de la universidad para instalar este sistema y elegir los algoritmos que den mejores resultados.

También se pretende identificar qué requisitos técnicos se necesitan para ponerlo en marcha y ver si se puede combinar con otras herramientas de seguridad. Además, se quiere evaluar qué tanto influye el uso de esta tecnología en la imagen que tiene la universidad, tanto en lo tecnológico como en lo relacionado con la seguridad.

2.2. MARCO TEÓRICO

2.2.1 Inteligencia Artificial

Desde una perspectiva descriptiva, el artículo de Navarro-Dolmestch (2023) explora el impacto potencial que las tecnologías de inteligencia artificial generativas (IAG) pueden tener sobre la integridad académica, centrándose en su influencia en la docencia y los procesos evaluativos en el ámbito universitario, especialmente en la enseñanza del derecho. La integridad académica es concebida como un conjunto de valores que se ven amenazados por la dependencia excesiva en la IAG. Entre los riesgos destacados se incluyen la posible inviabilidad del proyecto pedagógico y la disminución de la competitividad de las instituciones educativas. Para contrarrestar estas amenazas, se proponen cuatro medidas clave de mitigación dirigidas a preservar la integridad académica en los entornos universitarios.

2.2.2 Visión Artificial

Pinedo et al., (2021) explican que la visión artificial es una rama de la inteligencia artificial que trabaja con el procesamiento de imágenes para reconocer patrones, usando algoritmos en entornos controlados y repitiendo varias veces los procesos para afinar los resultados. Con tantos dispositivos que hoy en día pueden capturar imágenes (como cámaras y sensores), se está generando una gran cantidad de datos visuales que pueden ser súper útiles para que tanto instituciones públicas como privadas tomen mejores decisiones.

En su estudio, los autores se plantearon varios objetivos: mejorar el reconocimiento de patrones usando visión artificial, medir qué tan efectivo era ese sistema, implementarlo y analizar qué relación había entre esa mejora y el uso de la tecnología. La investigación fue cuasi-experimental, con un enfoque transversal, y se trabajó con una muestra de 8 patrones de imágenes.

Para probar el sistema, se usó una técnica de verificación con listas de chequeo aplicadas a dos grupos: uno de control y otro experimental. Los resultados fueron bastante claros: el grupo experimental procesó las imágenes en un promedio de 10,75 segundos, mientras que el grupo control tardó mucho más, unos 67,75 segundos. Además, se encontró una correlación del 72 % entre el reconocimiento de patrones y el uso del sistema de visión artificial, lo que demuestra que esta tecnología puede ser muy efectiva.

2.2.3 Visión Computacional

La visión computacional es una rama de la tecnología que busca enseñar a las máquinas a "ver" y entender lo que hay en imágenes o videos. A diferencia de nosotros, que procesamos lo que vemos casi sin pensar, las máquinas necesitan de algoritmos complejos que les ayuden a dividir las imágenes en partes clave para poder analizarlas.

Esta disciplina incluye varias técnicas avanzadas como segmentación de imágenes, detección de bordes y reconocimiento de patrones. Todo esto es muy útil en áreas como la robótica, el diagnóstico médico o los carros que se manejan solos. Para que funcione bien, la visión computacional necesita tanto buen hardware —como cámaras y sensores de calidad— como algoritmos que puedan interpretar la información visual con precisión, incluso en lugares con poca luz o en escenas donde hay muchos elementos que pueden confundir a la máquina (Szeliski, 2021).

En los últimos años, esta área ha avanzado muchísimo gracias a las redes neuronales convolucionales (CNN), que han logrado que el análisis de imágenes sea mucho más preciso y eficiente que antes.

2.2.4 Reconocimiento facial

El reconocimiento facial es una de las tecnologías más importantes dentro del área de visión computacional, ya que permite identificar o confirmar quién es una persona a partir de una foto o un video. Para hacerlo, usa algoritmos que analizan rasgos únicos del rostro, como la distancia entre los ojos, la forma de la nariz o la mandíbula. Con eso, crea algo así como una "huella facial" que es diferente en cada individuo.

Esta tecnología se está usando en muchas cosas prácticas, como en sistemas de seguridad, control de accesos o para desbloquear el celular con solo mirar la pantalla. Sin embargo, también ha generado algunas preocupaciones, sobre todo en lo que tiene que ver con la privacidad y el uso indebido de los datos biométricos, en especial cuando se aplica sin que la persona dé su consentimiento.

Gracias a los avances recientes, sobre todo con el uso del aprendizaje profundo, ahora estos sistemas pueden funcionar bien incluso cuando hay poca luz, cambios en la postura o diferentes expresiones en el rostro. Esto ha hecho que la precisión del reconocimiento facial mejore muchísimo (Sharma et al., 2020).

2.2.5 Algoritmo

Los algoritmos son el corazón de cualquier sistema de reconocimiento facial, y su desarrollo ha sido clave para que estos sistemas sean cada vez más precisos y rápidos. En un inicio, se usaban técnicas más tradicionales, como el Análisis de Componentes Principales (PCA), que básicamente reduce la cantidad de información en las imágenes de rostros para quedarse con lo más importante y así poder diferenciarlos. Aunque fue útil, esta técnica tiene sus fallas, sobre todo cuando cambian las condiciones de luz, la posición del rostro o las expresiones.

Después apareció el Análisis Discriminante Lineal (LDA), que mejoró un poco el rendimiento porque ayudaba a separar mejor los datos, es decir, a distinguir un rostro de otro con más claridad. Pero igual tenía problemas, especialmente cuando se trabajaba con pocos datos, ya que podía sobre ajustarse.

Hoy en día, el juego cambió completamente con la llegada de las Redes Neuronales Convolucionales (CNN). Estas redes aprenden por sí solas qué partes de una imagen son importantes, sin que haya que procesar tanto la imagen antes. Son capaces de manejar grandes cantidades de información y reconocer rostros en distintas condiciones, como con diferentes ángulos, luces o expresiones. Por eso, se han convertido en la opción más usada actualmente en sistemas de reconocimiento facial (Navarro-Dolmestch, 2023).

2.2.5.1 Tipos de algoritmos

- **Algoritmos de Búsqueda y Ordenación**

Los algoritmos de búsqueda y ordenación son fundamentales para el manejo eficiente de datos. Se utilizan para encontrar información específica dentro de grandes conjuntos de datos y para ordenar estos datos de manera que sean más fáciles de manejar.

  - Quicksort y Merge Sort son algoritmos de ordenación eficientes que utilizan la técnica de "divide y vencerás". Quicksort tiene un rendimiento promedio de O(n log n), mientras que Merge Sort garantiza un rendimiento consistente de O(n log n), siendo más eficiente en ciertos contextos (Cormen et al., 2022).
  - Búsqueda Binaria es un algoritmo de búsqueda rápida que encuentra elementos en una lista ordenada con una complejidad de O(log n),

permitiendo búsquedas rápidas en grandes conjuntos de datos (Sedgewick, 2020).

- **Algoritmos Greedy**

Los algoritmos greedy son aquellos que toman decisiones basadas en la mejor opción disponible en el momento, con la esperanza de encontrar una solución óptima o cercana a la óptima. Estos algoritmos son útiles en problemas de optimización.

  - Algoritmo de Dijkstra: Se utiliza para encontrar el camino más corto en un grafo, ideal para sistemas de navegación y redes de telecomunicaciones (Sedgewick, 2020).
  - En la computación distribuida, algoritmos greedy se utilizan para la optimización de recursos y asignación de tareas (Zhou et al., 2020).

- **Algoritmos de Programación Dinámica**

Los algoritmos de programación dinámica resuelven problemas complejos dividiéndolos en subproblemas más simples. Se utilizan en situaciones donde los subproblemas se repiten, lo que permite almacenar soluciones intermedias para reducir el tiempo de cálculo.

  - Fibonacci es un problema clásico que se resuelve eficientemente mediante programación dinámica, almacenando resultados previos en lugar de recalcularlos repetidamente. Aplicaciones modernas incluyen optimización en redes y análisis financiero (Goodrich & Tamassia, 2021).

- **Algoritmos Genéticos**

Los algoritmos genéticos son algoritmos de búsqueda y optimización que se basan en los principios de la evolución biológica. Utilizan técnicas como selección, cruce y mutación para iterar sobre posibles soluciones y optimizar resultados en problemas donde el espacio de búsqueda es muy grande.

  - Estos algoritmos son ampliamente utilizados en el diseño de redes neuronales, optimización de procesos y resolución de problemas NP-completos (Goldberg, 2021; Holland, 2020).

- **Algoritmos de Aprendizaje Automático**

Dentro del aprendizaje automático, los algoritmos supervisados y no supervisados permiten que las máquinas aprendan patrones a partir de datos.

- Redes Neuronales Convolucionales (CNN): Utilizadas principalmente para tareas de visión computacional, como la clasificación de imágenes y el reconocimiento de objetos. Las CNN son fundamentales para sistemas de visión autónoma, procesamiento de imágenes médicas y más (Goodfellow et al., 2016; LeCun et al., 2021).
- Máquinas de Soporte Vectorial (SVM): Utilizadas para la clasificación y regresión en conjuntos de datos etiquetados, se destacan por su robustez en problemas de alta dimensionalidad (Vapnik, 2021).

- **Algoritmos de Grafos**

Los algoritmos de grafos se utilizan para resolver problemas en estructuras complejas que involucran relaciones entre objetos, como redes sociales, mapas y sistemas de comunicación.

- Floyd-Warshall y Bellman-Ford son algoritmos clásicos para encontrar los caminos más cortos entre todos los pares de nodos en un grafo ponderado. Estos algoritmos se aplican en áreas como la optimización de redes de transporte y telecomunicaciones (Cormen et al., 2022; Floyd, 2020).

- **Algoritmos Recursivos**

Un algoritmo recursivo se llama a sí mismo para resolver subproblemas más pequeños de un problema mayor. Esto es particularmente útil en algoritmos de ordenación y búsqueda, y en problemas como la generación de fractales y árboles.

- Merge Sort: Divide repetidamente un conjunto de datos en mitades hasta que se alcanza una base trivial y luego las ordena de manera eficiente. Se utiliza en grandes bases de datos y sistemas distribuidos (Knuth, 2020).

2.2.6 Algoritmo de reconocimiento facial

El reconocimiento facial es una de las técnicas biométricas más usadas hoy en día, tanto en temas de seguridad como en redes sociales e incluso en los celulares para desbloquear la pantalla. Gracias a los avances en el aprendizaje profundo, los algoritmos que se usan han mejorado muchísimo, logrando ser más precisos y rápidos.

A continuación, se muestran los principales algoritmos que se utilizan actualmente para hacer reconocimiento facial.

- Eigenfaces y Fisherfaces

El método de Eigenfaces se basa en la descomposición de imágenes faciales mediante Análisis de Componentes Principales (PCA), reduciendo la dimensionalidad y destacando las características principales. Aunque fue revolucionario en los años 90, su desempeño en entornos con variaciones de iluminación y expresión es limitado (Cao et al., 2020). Fisherfaces mejora el enfoque de Eigenfaces al usar Análisis Discriminante Lineal (LDA), optimizando la separación entre diferentes clases de caras, lo que lo hace más robusto en estos entornos (Belhumeur et al., 1997).

- Redes Neuronales Convolucionales (CNN)

Las CNN se han convertido en el estándar para el reconocimiento facial moderno. Utilizan capas convolucionales para extraer características faciales jerárquicas. Modelos como FaceNet (Schroff et al., 2015) y VGGFace (Parkhi et al., 2015) alcanzan una precisión de nivel humano. Estos modelos proyectan imágenes faciales en un espacio de características donde las distancias euclidianas reflejan la similitud entre rostros. Las CNN también permiten la identificación a gran escala en bases de datos con millones de imágenes (Wang et al., 2021).

- DeepFace

DeepFace, desarrollado por Facebook, fue uno de los primeros sistemas en superar el 97% de precisión en la verificación facial, acercándose al rendimiento humano. Utiliza una red neuronal profunda para mapear imágenes faciales en un espacio tridimensional, abordando variaciones de pose, iluminación y expresión (Taigman et al., 2014). DeepFace es capaz de reconocer rostros en redes sociales y sistemas de autenticación.

- Multi-task Cascaded Convolutional Networks (MTCNN)

MTCNN es una red neuronal profunda que realiza detección y alineación facial simultáneamente. Esta arquitectura en cascada permite localizar puntos faciales clave como los ojos, nariz y boca, mejorando la precisión del reconocimiento facial en tiempo real (Zhang et al., 2016). MTCNN es ideal para tareas que requieren detectar caras en diferentes posiciones y tamaños dentro de una imagen.

- Modelos Basados en Aprendizaje Profundo

Los algoritmos más recientes están integrando arquitecturas más avanzadas como Transformers, que han demostrado ser útiles para el reconocimiento facial en entornos

no controlados. Estos modelos permiten el entrenamiento con grandes volúmenes de datos no etiquetados, mejorando la generalización y precisión en situaciones desafiantes, como cambios drásticos en la iluminación y la pose (Dosovitskiy et al., 2021).

- Consideraciones Éticas y Privacidad

El uso masivo del reconocimiento facial plantea desafíos éticos significativos, como la violación de la privacidad y el sesgo algorítmico. Investigaciones recientes han abordado estos temas, proponiendo regulaciones y mejoras técnicas para reducir el sesgo y garantizar que los sistemas respeten los derechos de los usuarios (Miller et al., 2021). Estos estudios sugieren que la transparencia y el uso ético son esenciales para la adopción a largo plazo del reconocimiento facial en la sociedad.

2.2.7 Reconocimiento Facial y Embeddings

El reconocimiento facial es una tecnología basada en la visión por computadora y el aprendizaje profundo (deep learning) que permite identificar o verificar la identidad de una persona a partir de sus rasgos faciales. Este proceso implica la detección del rostro, la extracción de características distintivas y la comparación de dichas características con una base de datos previamente registrada. Su aplicación se ha extendido ampliamente en entornos de seguridad, control de acceso y sistemas inteligentes de videovigilancia (Wang, Deng, & Hu, 2021).

En este contexto, los embeddings faciales representan el modelo matemático o vectorial generado por una red neuronal profunda que codifica los rasgos únicos del rostro humano en un conjunto de valores numéricos. Cada embedding actúa como una representación digital única del rostro, permitiendo calcular el nivel de similitud entre dos imágenes mediante métricas como la distancia euclidiana o la similitud coseno (Schroff, Kalenichenko, & Philbin, 2022).

2.2.8 Embeddings Faciales: Fundamentos Matemáticos y Aplicaciones

Los embeddings faciales constituyen representaciones vectoriales numéricas que transforman las características geométricas y texturales de un rostro humano en vectores de alta dimensionalidad dentro de un espacio matemático euclidiano. Estas representaciones permiten que la similitud facial se exprese como proximidad geométrica, facilitando comparaciones eficientes mediante cálculos de distancia euclidiana, coseno o métricas de Minkowski (Great Learning, 2024).

La arquitectura fundamental de los embeddings faciales mapea características faciales distintivas a espacios vectoriales de dimensiones típicamente entre 128 y 2048 elementos, donde cada dimensión codifica aspectos específicos de la morfología facial. Esta transformación debe mantener propiedades críticas como invarianza ante transformaciones fotométricas, robustez frente a variaciones de pose, y capacidad discriminante para distinguir identidades únicas mientras agrupa consistentemente múltiples imágenes de la misma persona (Schroff et al., 2024).

2.2.8.1 Propiedades Matemáticas de los Embeddings

Los embeddings faciales efectivos exhiben propiedades matemáticas específicas que garantizan su utilidad para reconocimiento: compacidad intra-clase, donde múltiples imágenes de la misma persona generan embeddings cercanos; separabilidad inter-clase, donde diferentes personas producen embeddings distantes; y estabilidad temporal, donde embeddings de la misma persona mantienen consistencia a lo largo del tiempo (Wang et al., 2022).

La función de mapeo f: I → R^d transforma una imagen facial I a un vector d-dimensional, donde la distancia euclidiana $||f(I_1) - f(I_2)||$ correlaciona inversamente con la similitud facial entre I_1 e I_2. Esta propiedad permite establecer umbrales de decisión para tareas de verificación e identificación facial.

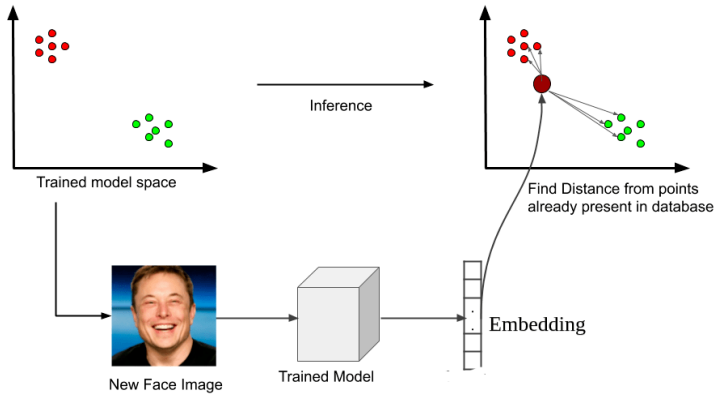


***Figura 1**. Espacio de embeddings faciales*

Fuente: Yash (2022). Face Recognition with FaceNet. Tech Musings.

2.2.9 Aprendizaje Profundo y Arquitecturas Neuronales Avanzadas

El aprendizaje profundo ha revolucionado el campo del reconocimiento facial mediante el desarrollo de arquitecturas neuronales especializadas capaces de aprender representaciones jerárquicas complejas directamente de datos raw. Las

redes neuronales profundas superan consistentemente a métodos tradicionales como PCA y LDA al capturar patrones no lineales y relaciones complejas en las estructuras faciales (Melzi et al., 2024).

2.2.9.1 Redes Neuronales Convolucionales (CNN) Especializadas

Las CNN implementan la operación de convolución matemática que permite la detección automática de características mediante filtros aprendibles. La operación convolucional se define como $(I * K)(i,j) = \Sigma_m \Sigma_n I(m,n)K(i-m,j-n)$, donde I representa la imagen de entrada y K el kernel convolucional (He et al., 2023). Esta operación preserva la estructura espacial de las imágenes mientras extrae características relevantes para reconocimiento.

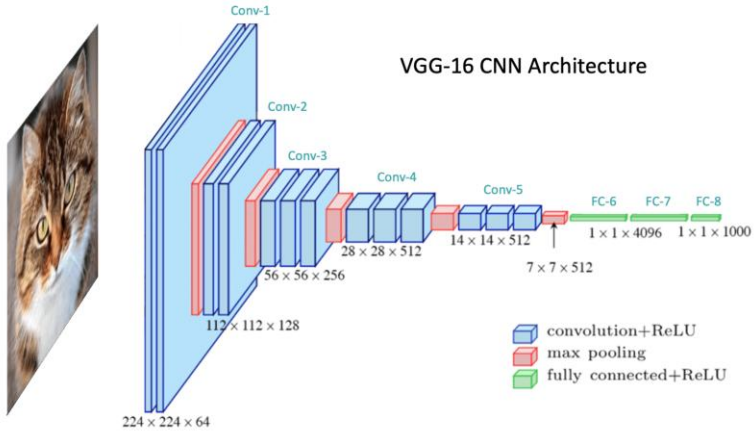


***Figura 2**. Arquitectura de una red neuronal (CNN)*

Fuente: LearnOpenCV (2023). "Convolutional Neural Network: A Complete Guide".

Las arquitecturas modernas incorporan componentes especializados como batch normalization para estabilizar el entrenamiento, dropout para regularización, y funciones de activación no lineales como ReLU y sus variantes. La profundidad de estas redes permite el aprendizaje de características desde bordes básicos en capas iniciales hasta representaciones semánticas complejas en capas finales (Tan & Le, 2023).

2.2.9.2 Arquitecturas Residuales y de Atención

Las redes residuales (ResNet) introducen conexiones skip que permiten el flujo directo de gradientes, facilitando el entrenamiento de redes extremadamente profundas. La función residual $F(x) = H(x) - x$ permite que las capas aprendan residuos en lugar de

mapeos completos, mitigando problemas de degradación y desvanecimiento del gradiente (He et al., 2023).

Los mecanismos de atención permiten que los modelos enfoquen selectivamente regiones faciales más informativas. La atención espacial asigna pesos diferentes a regiones de la imagen, mientras que la atención de canal pondera la importancia de diferentes mapas de características. Estos mecanismos mejoran significativamente la capacidad discriminante del modelo en presencia de oclusiones o variaciones de pose (Selvaraju et al., 2023).

2.2.9.3 Modelos Ligeros y Eficiencia Computacional

El desarrollo de modelos ligeros ha emergido como necesidad crítica para implementación en dispositivos con recursos limitados. MobileNets utilizan convoluciones separables en profundidad que factorizan convoluciones estándar en convoluciones depthwise y pointwise, reduciendo significativamente parámetros y operaciones computacionales (Howard et al., 2024).

ShuffleNet introduce operaciones de mezclado de canales que permiten el intercambio de información entre diferentes grupos de canales, manteniendo alta precisión con menor complejidad. EfficientNet optimiza simultáneamente profundidad, ancho, y resolución de red mediante escalamiento compuesto, logrando mejor trade-off entre precisión y eficiencia (Deng et al., 2023).

2.2.10 Algoritmos Especializados de Reconocimiento Facial

2.2.10.1 FaceNet y Aprendizaje de Espacios Métricos

FaceNet representa un paradigma transformador que aprende directamente una función de embedding que mapea imágenes faciales a espacios euclidianos donde distancias corresponden a similitudes faciales. La arquitectura utiliza pérdida triplet $L = \max(0, ||f(xa) - f(xp)||_2^2 - ||f(xa) - f(xn)||_2^2 + a)$, donde xa, xp, xn representan imágenes anchor, positiva, y negativa respectivamente, y a es el margen de separación (Karnati et al., 2023).

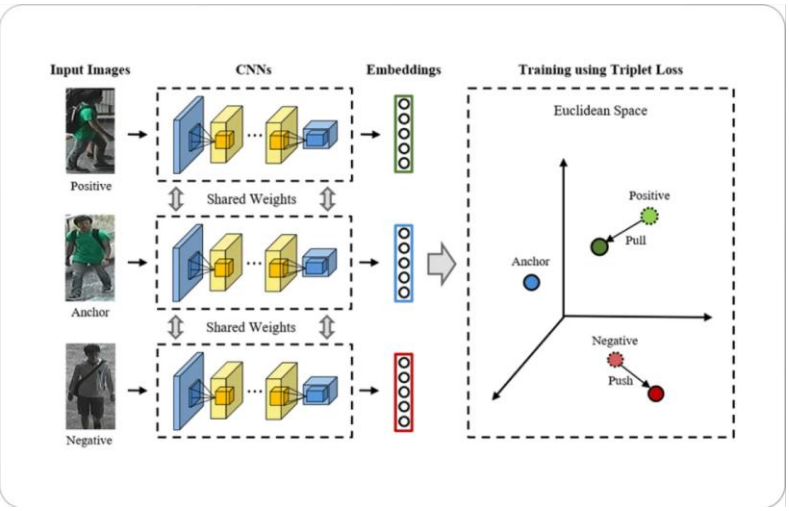


***Figura 3**.Visualización de la pérdida triplet.*

Fuente: V7 Labs (2023). Triplet loss: intro, implementation, use cases.

El entrenamiento con pérdida triplet optimiza simultáneamente la minimización de distancias intra-clase y maximización de distancias inter-clase. La selección de triplets es crítica para convergencia efectiva, utilizando estrategias como hard negative mining y semi-hard negative selection para identificar ejemplos más informativos durante el entrenamiento (Schroff et al., 2024).

2.2.10.2 ArcFace y Pérdidas Basadas en Margen

ArcFace introduce una función de pérdida que agrega margen angular en el espacio de características, mejorando la capacidad discriminante mediante L = -log(e^(s·cos(θyi+m))/(e^(s·cos(θyi+m)) + $\Sigma_{j\neq yi}$ e^(s·cosθj))), donde θyi es el ángulo entre la característica y el peso del centro de clase yi, s es el parámetro de escala, y m el margen angular (Wang et al., 2022).

Esta aproximación mejora la compacidad intra-clase y separabilidad inter-clase comparado con pérdidas tradicionales como softmax, resultando en embeddings más discriminantes y robustos para reconocimiento facial en condiciones desafiantes.

2.2.10.3 Multi-task CNN (MTCNN) para Detección y Alineación

MTCNN implementa una cascada de tres redes especializadas que realizan detección facial y localización de landmarks simultáneamente. P-Net genera propuestas iniciales de regiones candidatas, R-Net refina detecciones eliminando falsos positivos, y O-Net produce detecciones finales con landmarks faciales precisos (Zhang et al., 2023).

La arquitectura multi-tarea optimiza simultáneamente pérdidas de clasificación, regresión de bounding box, y localización de landmarks: L = αLcls + βLbox + γLlandmark, donde α, β, γ son factores de ponderación que balancean la contribución de cada tarea durante el entrenamiento.

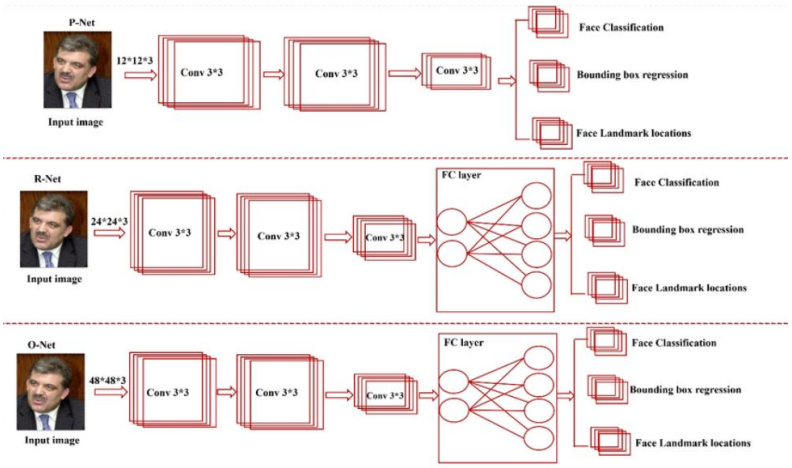


***Figura 4**. Arquitectura en cascada con sus etapas P-Net, R-Net y O-Net para detección y alineación facial.*

Fuente: LearnOpenCV (2023). "Convolutional Neural Network: A Complete Guide".

2.2.11 Preprocesamiento y Aumentación de Datos Avanzada

2.2.11.1 Normalización y Alineación Geométrica

El preprocesamiento constituye una etapa fundamental que estandariza las condiciones de entrada para garantizar robustez y consistencia del sistema. La normalización fotométrica mitiga variaciones de iluminación mediante técnicas como equalización de histograma, normalización de contraste local, y corrección gamma adaptativa (Analytics Vidhya, 2024).

La alineación geométrica utiliza landmarks faciales detectados para aplicar transformaciones afines que estandarizan pose y orientación facial. La transformación típica incluye rotación, escalado, y traslación para alinear ojos y boca en posiciones fijas, reduciendo variabilidad intra-clase y mejorando consistencia de embeddings.

2.2.11.2 Técnicas de Aumentación de Datos

La aumentación de datos expande artificialmente el dataset de entrenamiento mediante transformaciones que preservan identidad mientras introducen variabilidad realista. Las transformaciones incluyen rotaciones controladas (-15° a +15°), cambios

de escala (0.9x a 1.1x), traslaciones menores, y ajustes de brillo/contraste (Sandler et al., 2023).

Técnicas avanzadas incluyen elastic deformation para simular variaciones naturales de expresión facial, oclusiones sintéticas para mejorar robustez ante obstrucciones parciales, y color jittering para simular variaciones de iluminación y condiciones de captura diversas.

2.2.11.3 Generación de Datos Sintéticos

Los avances recientes en GANs (Generative Adversarial Networks) permiten la generación de rostros sintéticos fotorrealistas para aumentar datasets de entrenamiento. StyleGAN y sus variantes generan imágenes faciales de alta calidad con control sobre atributos específicos como edad, género, y expresión (Melzi et al., 2024).

La generación de datos sintéticos aborda problemas de sesgo demográfico y escasez de datos para grupos subrepresentados, permitiendo el entrenamiento de modelos más equitativos y robustos. Sin embargo, requiere cuidadosa validación para evitar artifacts que puedan afectar negativamente la generalización del modelo.

2.2.12 Métricas de Evaluación y Análisis de Performance

2.2.12.1 Métricas Biométricas Fundamentales

La evaluación rigurosa de sistemas de reconocimiento facial requiere métricas especializadas que reflejen performance en escenarios operacionales reales. La Tasa de Falsa Aceptación (FAR) cuantifica la probabilidad de que el sistema identifique incorrectamente a un impostor como usuario autorizado: FAR = FP/(FP + TN), donde FP son falsos positivos y TN verdaderos negativos (Liu et al., 2023).

La Tasa de Falso Rechazo (FRR) mide la probabilidad de que el sistema rechace incorrectamente a un usuario genuino: FRR = FN/(FN + TP), donde FN son falsos negativos y TP verdaderos positivos. El balance entre FAR y FRR determina la usabilidad y seguridad del sistema, requiriendo optimización según los requisitos específicos de la aplicación.

La Tasa de Error Igual (EER) representa el punto de operación donde FAR = FRR, proporcionando una métrica unificada para comparar sistemas. Un EER menor indica

mejor performance general, aunque la métrica óptima depende de los requisitos de seguridad versus usabilidad de la aplicación específica (Wang et al., 2022).

2.2.12.2 Curvas ROC y Métricas Derivadas

Las curvas ROC (Receiver Operating Characteristic) visualizan el trade-off entre la Tasa de Verdaderos Positivos (TPR = 1 - FRR) y la Tasa de Falsos Positivos (FPR = FAR) para diferentes umbrales de decisión. El Área Bajo la Curva (AUC) proporciona una medida agregada de performance independiente del umbral, donde AUC = 1 indica performance perfecta y AUC = 0.5 indica performance aleatoria (Selvaraju et al., 2023).

Las curvas DET (Detection Error Tradeoff) plotean FAR versus FRR en escala logarítmica, proporcionando mejor visualización de performance en regiones de baja tasa de error. Estas métricas son especialmente útiles para sistemas de alta seguridad donde tanto FAR como FRR deben ser minimizados.

2.2.12.3 Métricas de Performance Computacional

La evaluación de sistemas de videovigilancia requiere análisis detallado de métricas temporales y de throughput. La latencia end-to-end incluye tiempo de captura (tcapture), preprocesamiento (tpreprocess), inferencia del modelo (tinference), y post-procesamiento (tpost): Ltotal = tcapture + tpreprocess + tinference + tpost (Deng et al., 2023).

El throughput mide frames procesados por unidad de tiempo (fps), determinando la capacidad del sistema para manejar múltiples streams simultáneos. La eficiencia computacional se evalúa mediante métricas como FLOPS (Floating Point Operations Per Second) y consumo de memoria, críticas para implementación en dispositivos con recursos limitados.

2.2.12.4 Métricas de Robustez y Generalización

La robustez evalúa la estabilidad de performance ante variaciones en condiciones de operación. Las métricas incluyen degradación de accuracy bajo diferentes condiciones de iluminación, variaciones de pose (-90° a +90° para yaw, -60° a +60° para pitch y roll), y presencia de oclusiones (parciales hasta 50% del rostro) (Karnati et al., 2023).

La generalización cross-dataset evalúa performance cuando el modelo entrenado en un dataset se evalúa en datasets diferentes, indicando la capacidad del sistema

para funcionar en entornos no vistos durante entrenamiento. Esta métrica es crítica para sistemas desplegados en condiciones reales variables.

2.2.13 Infraestructura y Arquitectura de Sistemas Distribuidos

2.2.13.1 Arquitecturas de Microservicios

Los sistemas modernos de videovigilancia inteligente adoptan arquitecturas distribuidas basadas en microservicios que descomponen funcionalidades complejas en servicios independientes, cada uno responsable de una tarea específica. Esta aproximación facilita escalabilidad horizontal, mantenimiento independiente, y integración de nuevas capacidades sin afectar el sistema completo (Analytics Vidhya, 2024).

La arquitectura típica incluye servicios especializados: Video Ingestion Service para captura y streaming, Face Detection Service para localización de rostros, Feature Extraction Service para generación de embeddings, Identity Matching Service para comparación con base de datos, y Alert Management Service para generación y distribución de notificaciones.

2.2.13.2 Protocolos de Comunicación Inter-Servicios

La comunicación entre microservicios utiliza protocolos ligeros y standards de la industria. REST APIs proporcionan interfaces síncronas para consultas directas, mientras que message queues (RabbitMQ, Apache Kafka) facilitan comunicación asíncrona para procesamiento de alto volumen (Tan & Le, 2023).

gRPC emerge como protocolo eficiente para comunicación entre servicios, utilizando Protocol Buffers para serialización compacta y soporte nativo para streaming bidireccional. Esta tecnología reduce latencia y overhead comparado con REST tradicional, especialmente crítico para sistemas de tiempo real.

2.2.13.3 Load Balancing y Alta Disponibilidad

Los sistemas de videovigilancia requieren alta disponibilidad y capacidad de manejar cargas variables. Load balancers distribuyen requests entre múltiples instancias de servicios, utilizando algoritmos como round-robin, least connections, o weighted distribution según características específicas de cada servicio (Howard et al., 2024).

La implementación incluye health checks automáticos para detectar instancias no disponibles, circuit breakers para prevenir propagación de fallos, y auto-scaling para ajustar dinámicamente el número de instancias según la demanda. La redundancia

geográfica mediante deployment en múltiples zonas de disponibilidad garantiza continuidad operacional ante fallos de infraestructura.

2.2.14 Edge Computing y Optimización de Modelos

2.2.14.1 Paradigmas de Procesamiento Distribuido

El edge computing emerge como paradigma crítico para sistemas de videovigilancia, procesando datos cerca del punto de captura para reducir latencia, minimizar ancho de banda, y mejorar privacidad. La arquitectura distribuye inteligencia entre edge devices, fog nodes, y cloud infrastructure según los requisitos específicos de cada aplicación (Deng et al., 2023).

Edge devices ejecutan modelos optimizados para detección facial básica y filtrado de eventos relevantes, reduciendo tráfico de red. Fog nodes realizan procesamiento intermedio como extracción de embeddings y correlación temporal. Cloud infrastructure maneja análisis complejos, almacenamiento a largo plazo, y entrenamiento de modelos.

2.2.14.2 Técnicas de Optimización de Modelos

La implementación en edge devices requiere optimización agresiva de modelos deep learning para cumplir restricciones de recursos. La quantización reduce precisión numérica de pesos y activaciones de float32 a int8, reduciendo tamaño de modelo hasta 4x con degradación mínima de accuracy (Sandler et al., 2023).

Model pruning elimina conexiones neuronales con pesos pequeños, creando modelos sparse que requieren menos memoria y computación. Structured pruning remueve canales o filtros completos, facilitando aceleración en hardware estándar. Unstructured pruning remueve pesos individuales, logrando mayor compresión pero requiriendo hardware especializado para aceleración efectiva.

La destilación de conocimiento transfiere conocimiento de modelos complejos (teacher) a modelos simples (student), manteniendo performance mientras reduce complejidad. El proceso optimiza L = aLCE(ys, y) + (1-a)LKD(ys, yt), donde ys, yt, y son salidas del student, teacher, y etiquetas verdaderas respectivamente (He et al., 2023).

2.2.14.3 Aceleración por Hardware

Los sistemas modernos utilizan hardware especializado para acelerar inferencia de deep learning. GPUs proporcionan paralelización masiva ideal para operaciones

matriciales, mientras que TPUs (Tensor Processing Units) están optimizadas específicamente para operaciones de machine learning (Tan & Le, 2023).

FPGAs (Field-Programmable Gate Arrays) ofrecen flexibilidad para implementar arquitecturas personalizadas con mayor eficiencia energética que GPUs. Los procesadores neuromorphic como Intel Loihi imitan estructuras neuronales biológicas, proporcionando eficiencia extrema para aplicaciones de inferencia.

2.2.15 Consideraciones Éticas, Legales y de Privacidad

2.2.15.1 Marco Regulatorio y Protección de Datos

La implementación de sistemas de reconocimiento facial en entornos académicos enfrenta marcos regulatorios complejos que varían según jurisdicción. En Ecuador, la Ley Orgánica de Protección de Datos Personales establece principios fundamentales para tratamiento de datos biométricos, incluyendo consentimiento informado, finalidad específica, y proporcionalidad (Karnati et al., 2023).

Los datos biométricos son clasificados como datos sensibles que requieren medidas de protección especiales. La implementación debe incorporar principios de privacy by design, implementando protecciones de privacidad desde el diseño inicial del sistema rather que como adición posterior.

2.2.15.2 Riesgos de Privacidad y Ataques de Reconstrucción

Los embeddings faciales, aunque diseñados para privacidad, presentan vulnerabilidades ante ataques sofisticados. Los ataques de reconstrucción intentan regenerar imágenes faciales a partir de embeddings almacenados, utilizando técnicas como gradient descent optimization o GANs inversas (IronCore Labs, 2024).

La mitigación requiere técnicas como differential privacy, agregando ruido gaussiano calibrado a embeddings: ε-differential privacy garantiza que la presencia o ausencia de un individuo en el dataset no puede ser determinada con alta confianza. Template protection transforma embeddings mediante funciones irreversibles que mantienen utilidad para comparación mientras previenen reconstrucción directa.

2.2.15.3 Sesgo Algorítmico y Equidad

Los sistemas de reconocimiento facial exhiben sesgos sistemáticos que afectan desproporcionalmente grupos demográficos específicos. Estudios documentan disparidades significativas en accuracy entre géneros, etnias, y grupos etarios, con particular degradación para mujeres de piel oscura (Melzi et al., 2024).

La medición de fairness utiliza métricas como Demographic Parity ($P(\hat{Y}=1 | A=0) = P(\hat{Y}=1 | A=1)$), Equalized Odds, y Calibration across grupos protegidos. La mitigación incluye diversificación de datasets de entrenamiento, re-weighting de muestras subrepresentadas, y adversarial debiasing durante entrenamiento.

2.2.15.4 Transparencia y Explicabilidad

La explicabilidad en sistemas de reconocimiento facial requiere técnicas para visualizar y entender decisiones del modelo. Grad-CAM (Gradient-weighted Class Activation Mapping) genera mapas de calor que destacan regiones faciales más influyentes en la decisión del modelo (Selvaraju et al., 2023).

LIME (Local Interpretable Model-agnostic Explanations) y SHAP (SHapley Additive exPlanations) proporcionan explicaciones locales para decisiones individuales, identificando características faciales específicas que contribuyen a identificaciones correctas o erróneas.

2.2.16 Protocolos de Comunicación y Streaming Multimedia

2.2.16.1 Protocolos de Streaming de Video

Los sistemas de videovigilancia requieren protocolos robustos para transmisión de video en tiempo real con mínima latencia y máxima calidad. RTSP (Real Time Streaming Protocol) constituye el standard para control de streaming multimedia, proporcionando comandos para play, pause, seeking, y control de calidad de stream (Analytics Vidhya, 2024).

RTMP (Real-Time Messaging Protocol) optimiza transmisión de baja latencia mediante chunks pequeños y buffering mínimo, ideal para aplicaciones interactivas. HTTP Live Streaming (HLS) implementa streaming adaptativo que ajusta automáticamente calidad según ancho de banda disponible, garantizando reproducción continua en condiciones de red variables.

2.2.16.2 Codificación y Compresión de Video

La eficiencia de codificación es crítica para sistemas de videovigilancia que manejan múltiples streams simultáneos. H.264/AVC utiliza predicción temporal inter-frame y espacial intra-frame, transformadas DCT, y codificación entrópica para lograr compresión eficiente manteniendo calidad visual (Zhang et al., 2023).

H.265/HEVC mejora eficiencia hasta 50% comparado con H.264 mediante unidades de codificación de tamaño variable, mejor predicción, y transforms adaptativos. AV1

es el codec más reciente que proporciona eficiencia superior especialmente para contenido de ultra alta resolución.

2.2.16.3 Calidad de Servicio (QoS) y Adaptación de Red

Los sistemas de videovigilancia implementan mecanismos de QoS para garantizar performance consistente ante variaciones de red. Traffic shaping prioriza tráfico de video crítico, mientras que adaptive bitrate streaming ajusta calidad automáticamente según condiciones de red (Howard et al., 2024).

Buffer management y jitter control minimizan latencia end-to-end crítica para respuesta en tiempo real a eventos de seguridad. Los protocolos incluyen mecanismos de recuperación de errores como Forward Error Correction (FEC) y Automatic Repeat reQuest (ARQ) para mantener calidad de stream ante pérdida de paquetes.

2.2.17 Gestión Avanzada de Bases de Datos

2.2.17.1 Bases de Datos Vectoriales Especializadas

El almacenamiento y búsqueda eficiente de embeddings faciales requiere bases de datos especializadas en datos vectoriales de alta dimensionalidad. Sistemas como Pinecone, Weaviate, y Chroma implementan índices optimizados para búsquedas de similaridad en espacios euclidianos complejos (Great Learning, 2024).

Los índices aproximados como Hierarchical Navigable Small World (HNSW) y Locality Sensitive Hashing (LSH) permiten búsquedas sub-lineales en datasets masivos, reduciendo complejidad de O(n) a O(log n) para búsquedas de nearest neighbor. Estos algoritmos balancean precisión con velocidad según los requisitos específicos de la aplicación.

2.2.17.2 Optimización de Consultas y Performance

La optimización de consultas para bases de datos vectoriales incluye técnicas como index warming para cargar índices en memoria, query result caching para consultas frecuentes, y batch processing para múltiples consultas simultáneas. El partitioning horizontal distribuye embeddings across múltiples nodos según criterios como user ID o timestamp (Analytics Vidhya, 2024).

III. METODOLOGÍA

3.1. ENFOQUE METODOLÓGICO

3.1.1. Enfoque Mixto

Según Arenas (2021), este método que se propone es bastante innovador porque, aunque toma ideas de los enfoques cualitativo (CUAL) y cuantitativo (CUAN), en realidad arma su propia forma de trabajar. Crea su propio marco teórico, su forma de diseñar y también cómo va a recolectar, procesar y analizar los datos.

La investigación se va a realizar con un enfoque mixto, es decir, usando tanto lo cualitativo como lo cuantitativo:

- En la parte **cualitativa**, se va a describir cómo son los dispositivos tecnológicos y también se van a comparar distintos algoritmos que se usan en el prototipo de reconocimiento facial.
- En cuanto al enfoque **cuantitativo**, se usará para medir qué tanto ha influido el sistema en la seguridad del lugar. Se van a contar cuántas pruebas se han hecho y cuántas personas hay registradas en la base de datos del sistema.

3.1.2. Tipo de Investigación

Investigación Exploratoria

La investigación exploratoria se puede ver como una especie de brújula que nos ayuda a orientarnos cuando estamos tratando temas nuevos o que todavía no han sido muy estudiados. Como dice Bautista (2022), su principal objetivo es descubrir, entender y empezar a conocer mejor un fenómeno del que no se tiene mucha información.

En este proyecto, este tipo de investigación va a ser súper útil en las primeras etapas, ya que nos va a permitir familiarizarnos con los conceptos básicos sobre los algoritmos de reconocimiento facial y también conocer las características técnicas de los equipos que vamos a usar.

Esta idea es clave para arrancar bien el proyecto, ya que nos ayuda a entender desde cero cómo funcionan estos algoritmos y qué tipo de tecnología necesitamos manejar para desarrollarlos correctamente.

Investigación – Acción

Según Flores (2021), la investigación-acción es una forma de trabajo que combina la teoría con la práctica, y además incluye la participación de las personas y mucha reflexión. En este proyecto se va a aplicar este enfoque para poder diseñar e implementar un algoritmo de reconocimiento facial dentro de un sistema de cámaras de seguridad.

La idea es que esta investigación ayude a resolver problemas concretos de seguridad que existen en el laboratorio, y para eso se va a contar con la participación de los propios usuarios. Así, el sistema que se desarrolle podrá ajustarse mejor a lo que realmente se necesita en ese entorno específico.

3.2. IDEA A DEFENDER

El uso de reconocimiento facial en los sistemas de video vigilancia fortalecerá la gestión de la seguridad física en la carrera de Computación de la Universidad Politécnica Estatal del Carchi.

3.3. DEFINICIÓN Y OPERACIONALIZACIÓN DE LAS VARIABLES

Tabla 1. Operacionalización de variables

Variable	Dimensión	Indicadores	Técnica	Instrumento
Independiente: Reconocimiento facial	Precisión de identificación	- Tasa de éxito en la identificación de rostros - Porcentaje de coincidencias correctas con base de datos - Porcentaje de rostros no reconocidos	Pruebas con base de datos de rostros	Software de análisis facial (FaceNet + Shinobi)
	Tasa de falsos positivos	Casos donde se identifica incorrectamente a un individuo	Revisión de logs y eventos	Bitácora de Shinobi / registros de predicción
	Tiempo de respuesta	- Tiempo promedio desde la captura de imagen hasta la identificación facial - Latencia del modelo durante el reconocimiento	Pruebas reales	Consola de sistema
	Estabilidad del sistema	- Cantidad de cámaras activas vs inactivas - Tiempos de caída del servidor - Disponibilidad del servicio	Monitoreo del sistema	Logs de servidor Ubuntu / panel de Shinobi
Dependiente: Videovigilancia	Detección y monitoreo facial	- Número total de rostros detectados - Registro de rostros desconocidos	Observación del sistema	Datos del Reconocimiento del sistema
	Integración con infraestructura	- Funcionalidad con cámaras IP ya existentes - Flujo estable de video en red - Uso eficiente de recursos del servidor	Verificación técnica	Análisis de conectividad y consumo del servidor

3.4. MÉTODOS UTILIZADOS

En esta investigación se usaron varios métodos, pero el que mejor se ajustó a lo que se quería lograr fue el método científico. Este enfoque es bastante completo, ya que permite recolectar información nueva de forma ordenada a través de la observación, la experimentación y el análisis de resultados. Además, es flexible, porque se pueden hacer ajustes según lo que se vaya descubriendo o lo que se necesite para comprobar la hipótesis.

Durante todo el proceso surgieron muchas preguntas, y eso fue súper útil porque ayudó a avanzar y mejorar el proceso de reconocimiento. Como explican Pérez y Gómez (2022), el método científico se basa en principios como la observación bien planificada, mediciones exactas, pruebas experimentales, análisis crítico de los datos y la constante revisión de hipótesis. Todo esto asegura que los resultados sean válidos y confiables.

También se hizo una observación detallada de cómo funcionaba el sistema, lo cual fue clave para asegurarse de que todo estuviera bien desarrollado y cumpliera con los objetivos.

3.5. ANÁLISIS ESTADÍSTICO

3.5.1.1Población y muestra

La población está comprendida por los estudiantes de la carrera de computación de la UPEC, con un total de **329**. Se realizó el muestreo de tipo probabilístico con población finita. Para ello, se empleó la fórmula propuesta por Arias (2016), la cual se expresa de la siguiente manera:

$$n = \frac{N * Z^2 * p * q}{e^2(N-1) + Z^2 * p * q}$$

Donde:

- **n**: Representa el tamaño de la muestra.
- **N**: Representa el total de integrantes de la población.
- **Z²**: Valor de confianza.
- **e**: Error muestral, normalmente entre 1% y 5%.
- **p**: Proporción de individuos con la característica investigada.
- **q**: Proporción complementaria de individuos sin la característica investigada.

En este caso se consideró un nivel de confianza del **95%** y un margen de error del **5%**. Con estos valores, el cálculo se realizó de la siguiente manera:

$$n = \frac{329 * (1.96)^2 * 0.5 * 0.5}{(0.05)^2(329 - 1) + (1.96)^2 * 0.5 * 0.5}$$

$$n = 176.9966 \approx 177$$

Por lo tanto, el tamaño de la muestra corresponde a **177 estudiantes**.

3.5.1.2. Técnicas e Instrumentos

Para recolectar los datos necesarios en esta investigación, se usaron dos herramientas principales: una encuesta estructurada y una observación sistemática. La encuesta sirvió para conocer lo que piensan los estudiantes sobre el tema, mientras que la observación ayudó a revisar de manera ordenada cómo funcionaba el prototipo en acción. Ambas técnicas fueron clave para sacar información útil que luego se analizó.

Encuesta

Se aplicó un cuestionario con preguntas cerradas dirigido especialmente a los estudiantes de la carrera de Computación. La idea principal era entender cómo ven ellos la implementación de un sistema de reconocimiento facial conectado a un circuito cerrado de cámaras. Este sistema está pensado para mejorar la seguridad en conjunto con el personal de seguridad que posee la carrera, buscando crear un ambiente más seguro y eficiente en los laboratorios y espacios tecnológicos que los estudiantes usan con frecuencia.

IV. RESULTADOS Y DISCUSIÓN

4.1. RESULTADOS

Después de realizar encuestas a estudiantes de la Carrera de Computación, se presentan los resultados de cada pregunta de la encuesta. Estos resultados muestran la valoración y opiniones de los estudiantes sobre el sistema que se va a desarrollar.

Pregunta 1 - En una escala del 1 al 5, donde 1 es "muy insatisfecho/a" y 5 es "muy satisfecho/a", ¿qué tan satisfecho/a está con el nivel de seguridad que ofrece el sistema de videovigilancia actual?

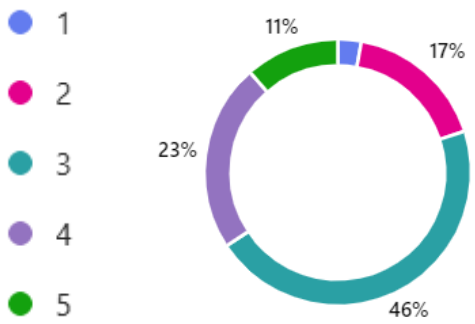


Figura 5 *Nivel de seguridad*

Descripción: En el grafico se muestra con una escala del 1 al 5, donde el 1 es "muy insatisfecho/a" y 5 es "muy insatisfecho/a" en cuanto a que satisfacción se encuentra con el nivel de seguridad que ofrece el sistema de vigilancia actual.

Análisis: La mayoría de los estudiantes, un 46%, se encuentra en una posición neutral respecto a su satisfacción con el sistema de seguridad. Le sigue un 23% que se siente *casi satisfecho/a* y un 17% que se encuentra *poco satisfecho/a*. Esto indica que, en general, los estudiantes tienen una percepción moderada del sistema de vigilancia, con una tendencia hacia la insatisfacción leve o la indiferencia.

Pregunta 2 - ¿Cree que el sistema de videovigilancia actual es efectivo para prevenir accesos no autorizados?

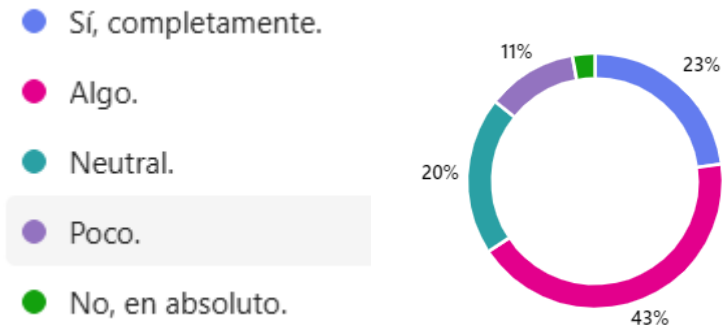


Figura 6. *Sistema Efectivo -Accesos no autorizados*

Descripción: En el grafico se muestra si los estudiantes creen que el sistema de videovigilancia actual es efectivo para prevenir accesos no autorizados en la carrera de computación.

Análisis: La mayoría de los estudiantes, un 43%, considera que el sistema de videovigilancia es *algo efectivo*. En segundo lugar, un 23% está completamente de acuerdo en que el sistema es efectivo, mientras que un 20% se mantiene neutral respecto a su eficacia. Estos resultados indican que, en general, los estudiantes perciben el sistema como moderadamente efectivo.

Pregunta 3 - ¿Ha notado incidentes de accesos no autorizados o brechas de seguridad en las instalaciones?

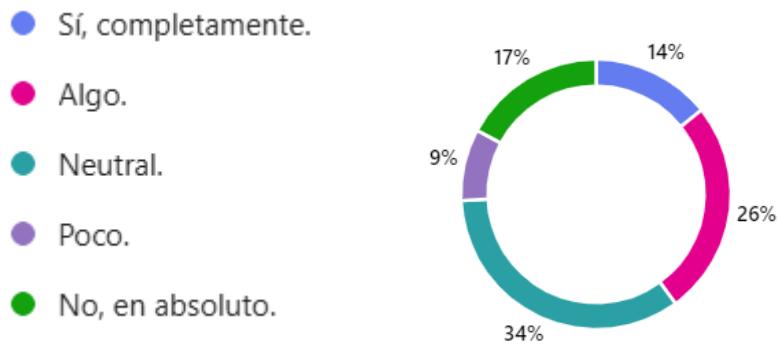


Figura 7*. Incidentes de accesos*

Descripción: La pregunta planteada a los estudiantes fue: "¿Ha notado incidentes de accesos no autorizados o brechas de seguridad en las instalaciones?". Los resultados del gráfico reflejan la percepción de los encuestados ante esta problemática.

Análisis: La mayor parte de los estudiantes (34%) se mostró neutral respecto a la existencia de incidentes de accesos no autorizados o brechas de seguridad. Un 26% indicó que ha percibido estas situaciones "en algo", mientras que el 14% afirmó haberlas notado "completamente". Por otro lado, el 17% señaló que apenas ha notado estas brechas y el 9% indicó no haberlas percibido en absoluto. Estos datos revelan una diversidad de opiniones, donde la neutralidad y la percepción moderada predominan, sugiriendo posibles dudas o falta de claridad sobre la seguridad actual.

Pregunta 4 - En caso de situaciones de emergencia o incidentes, ¿cómo describiría la respuesta del sistema actual?

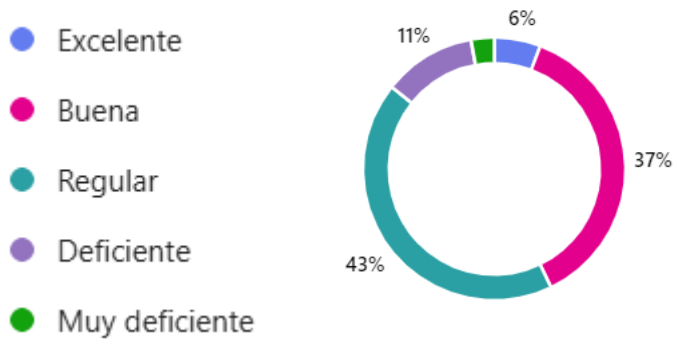


Figura 8. *Respuesta sistema actual*

Descripción: La pregunta planteada fue: "En caso de situaciones de emergencia o incidentes, ¿cómo describiría la respuesta del sistema actual?". El gráfico refleja las opiniones de los estudiantes sobre la efectividad del sistema en estas situaciones.

Análisis: La mayoría de los estudiantes, un 43%, considera que la respuesta del sistema es "regular", mientras que un 37% la evalúa como "buena". Por otro lado, el 11% califica la respuesta como "deficiente" y un 3% la describe como "muy deficiente". Solo el 6% de los encuestados percibe la respuesta como "excelente". Estos resultados indican que, aunque una parte significativa tiene una percepción positiva, existe una proporción relevante que identifica aspectos por mejorar en el sistema actual.

Pregunta 5 - ¿Qué tan de acuerdo está con la implementación de un sistema de reconocimiento facial como complemento al sistema de videovigilancia?

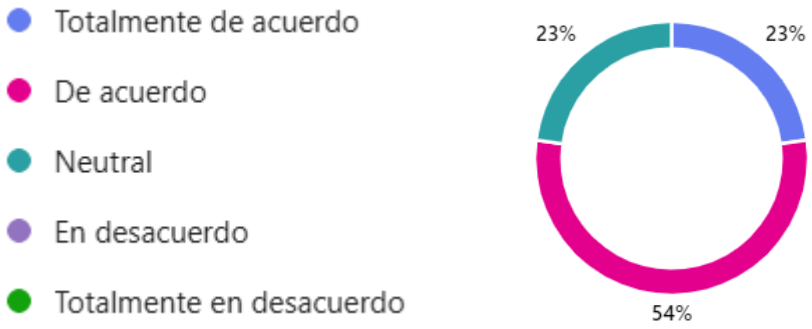


Figura 9. *Implementación Sistema*

Descripción: La pregunta formulada fue: "¿Qué tan de acuerdo está con la implementación de un sistema de reconocimiento facial como complemento al sistema de videovigilancia?". Los resultados muestran las posturas de los estudiantes respecto a esta propuesta.

Análisis: La mayoría de los estudiantes, un 54%, está "de acuerdo" con la implementación del sistema de reconocimiento facial, mientras que un 23% se muestra "totalmente de acuerdo". El 23% restante mantiene una postura "neutral". No se registraron respuestas en "desacuerdo" ni "totalmente en desacuerdo". Estos datos reflejan un alto nivel de aceptación entre los estudiantes hacia esta tecnología, con una tendencia clara hacia posturas positivas frente a su incorporación.

Pregunta 6 - ¿Qué nivel de comodidad siente al ser identificado/a mediante reconocimiento facial? (Escala Likert del 1 al 5: 1 = Muy incómodo/a, 5 = Muy cómodo/a?

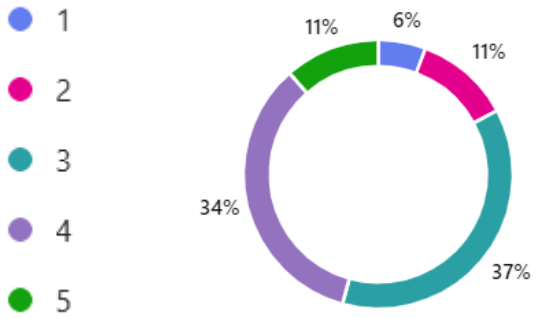


Figura 10. *Comodidad – Reconocimiento Facial*

Descripción: La pregunta realizada fue: "*¿Qué nivel de comodidad siente al ser identificado/a mediante reconocimiento facial?*". Los resultados se expresan en una escala Likert del 1 al 5, donde 1 es "Muy incómodo/a" y 5 es "Muy cómodo/a".

Análisis: El nivel de comodidad más frecuente fue el 3, seleccionado por el 37% de los estudiantes, lo que indica una postura neutral. El 34% manifestó comodidad, distribuyéndose entre el nivel 4 (34%) y el nivel 5 (11%). Por otro lado, un 17% expresó incomodidad, con el nivel 2 (6%) y el nivel 1 (11%). Estos resultados muestran que, aunque la mayoría de los estudiantes tiene una percepción neutral o positiva respecto al reconocimiento facial, un segmento menor siente incomodidad con esta tecnología.

Pregunta 7 - Según su opinión, ¿de qué manera el reconocimiento facial podría mejorar la identificación de personas autorizadas?

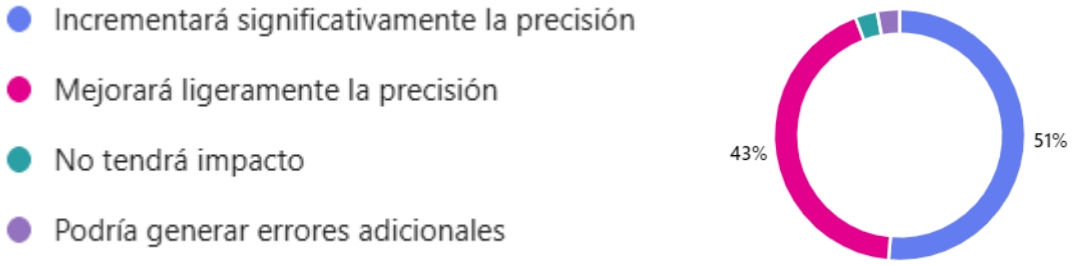


Figura 11. *Mejorar la identificación*

Descripción: La pregunta planteada fue: "*Según su opinión, ¿de qué manera el reconocimiento facial podría mejorar la identificación de personas autorizadas?*". Los resultados reflejan las expectativas de los estudiantes frente al impacto de esta tecnología.

Análisis: La mayoría de los estudiantes, un 51%, considera que el reconocimiento facial *incrementará significativamente la precisión* en la identificación de personas autorizadas. Un 43% opina que *mejorará ligeramente la precisión*. Por otro lado, solo un 3% piensa que *no tendrá impacto*, y otro 3% cree que *podría generar errores adicionales*. Estos resultados reflejan una percepción mayoritariamente positiva, con altas expectativas sobre el potencial de esta tecnología para optimizar la seguridad.

Pregunta 8 - ¿Qué tan confiable considera que sería un sistema de reconocimiento facial en términos de evitar errores de identificación?

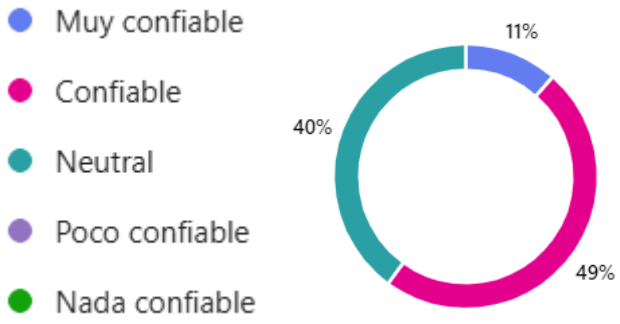


Figura 12. *Confianza del sistema*

Descripción: La pregunta realizada fue: "*¿Qué tan confiable considera que sería un sistema de reconocimiento facial en términos de evitar errores de identificación?*". Los resultados muestran la percepción de los estudiantes sobre la fiabilidad de esta tecnología.

Análisis: La mayoría de los estudiantes, un 49%, considera que el sistema sería *confiable*, mientras que un 11% lo califica como *muy confiable*. Un 40% se mantiene *neutral* respecto a su confiabilidad. No se registraron respuestas en las categorías de "poco confiable" ni "nada confiable". Estos resultados sugieren que los estudiantes tienen una percepción generalmente positiva sobre la confiabilidad del reconocimiento facial, aunque una parte significativa se mantiene cautelosa o no está completamente convencida.

Pregunta 9 - Desde su perspectiva, ¿qué impacto podría tener el reconocimiento facial en la seguridad de las instalaciones?

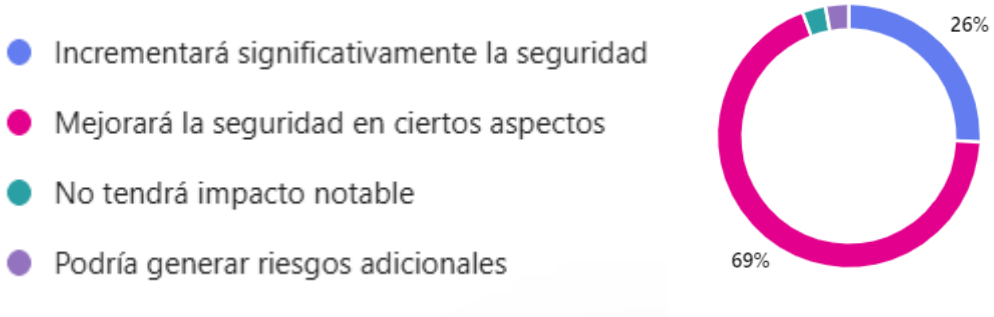


Figura 13. *Impacto – Reconocimiento*

Descripción: La pregunta formulada fue: "*Desde su perspectiva, ¿qué impacto podría tener el reconocimiento facial en la seguridad de las instalaciones?*". Los resultados muestran cómo los estudiantes perciben el impacto de esta tecnología en la seguridad.

Análisis: La mayoría de los estudiantes, un 69%, considera que el reconocimiento facial *mejorará la seguridad en ciertos aspectos*. Un 26% opina que *incrementará significativamente la seguridad*. Solo un 3% cree que *no tendrá impacto notable*, y otro 3% piensa que *podría generar riesgos adicionales*. Estos resultados reflejan una visión mayoritariamente positiva, con la mayoría de los estudiantes confiando en que el reconocimiento facial contribuirá a mejorar la seguridad, aunque algunos permanecen cautelosos ante posibles riesgos.

Pregunta 10 - ¿Cree que el uso de reconocimiento facial podría generar mayor confianza en la seguridad del lugar?

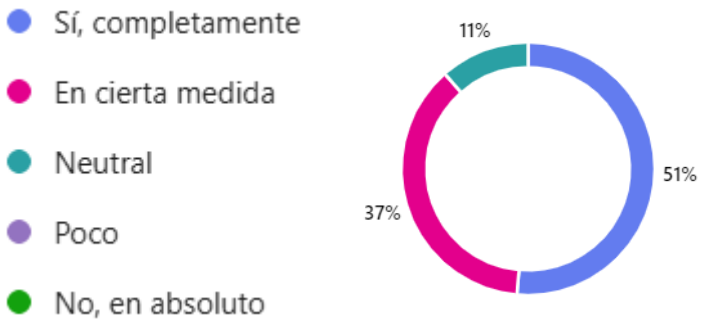


Figura 14. *Confianza en la seguridad*

Descripción: La pregunta planteada fue: "*¿Cree que el uso de reconocimiento facial podría generar mayor confianza en la seguridad del lugar?*". Los resultados muestran las opiniones de los estudiantes sobre la influencia del reconocimiento facial en la percepción de seguridad.

Análisis: La mayoría de los estudiantes, un 51%, cree que el uso de reconocimiento facial *generará mayor confianza completamente*. Un 37% opina que *lo hará en cierta medida*. Un 11% se mantiene *neutral* sobre su impacto. No se registraron respuestas en las categorías de "poco" ni "no, en absoluto". Estos resultados reflejan una percepción mayoritariamente positiva, indicando que la mayoría de los estudiantes confía en que el reconocimiento facial contribuirá a una mayor seguridad en el lugar.

4.1.1 Metodología Top -Down

Figura 15. *Metodología Top – Down*

4.1.1.1 Fase 1 – Recolección y Refinamiento de Requisitos

Después de realizar un análisis detallado del entorno de seguridad en la Universidad Politécnica Estatal del Carchi, se identificaron diversas problemáticas en el sistema de videovigilancia, específicamente en la detección e identificación de personas mediante reconocimiento facial.

El principal objetivo del proyecto es mejorar la eficiencia en la identificación de individuos en el circuito cerrado de videovigilancia, permitiendo una detección rápida y precisa de personas en tiempo real, con el fin de optimizar la seguridad dentro del campus universitario.

4.1.1.2. Datos Iniciales

Este proyecto tiene como objetivo principal mejorar la seguridad dentro de los espacios académicos usando un sistema de videovigilancia que incorpora reconocimiento facial. Para ello, se están utilizando cámaras Hikvision DS-2CD2020F-I, las cuales están conectadas a un sistema que analiza en tiempo real las imágenes captadas. El sistema es capaz de detectar rostros y generar alertas si se identifica algo fuera de lo normal.

Todo el procesamiento de video y la parte de inteligencia artificial se lleva a cabo en una laptop que cuenta con un procesador Intel Core i5 de 12ª generación y una tarjeta gráfica NVIDIA RTX 3050. Durante el desarrollo del proyecto, también se va a

analizar si este equipo es suficiente para manejar toda la carga del sistema o si será necesario usar un servidor adicional que apoye el procesamiento.

En esta primera etapa, el objetivo fue investigar y definir los requerimientos técnicos del sistema de reconocimiento facial enfocado en videovigilancia. Para lograr eso, se realizó una revisión bibliográfica y técnica sobre los algoritmos de reconocimiento facial, así como sobre el hardware y software necesarios y las mejores estrategias para procesar imágenes dentro de un entorno de videovigilancia.

4.1.1.3. Requisitos de Hardware

La configuración del hardware debe permitir el análisis eficiente de las imágenes en tiempo real.

Tabla 2. Requisitos de Hardware

Especificación	Propiedades
Cámara de videovigilancia	Hikvision DS-2CD2020F-I (1920×1080, H.264, IR)
Computador Principal	Servidor HPE ProLiant DL360 Gen10, Xeon Silver 4110, 125.5 GB RAM, 2.4 TB SSD, Ubuntu 22.04.3
Unidad de almacenamiento externo (opcional)	HDD 2TB o SSD externo para almacenamiento de grabaciones
Switch de red (si aplica)	Switch Gigabit para conexión de múltiples cámaras

- **Cámaras Hikvision DS-2CD2020F-I**

La cámara Hikvision DS-2CD2020F-I es un modelo de videovigilancia que destaca por su resolución Full HD de 1920×1080 píxeles, lo que permite obtener imágenes claras y con buen nivel de detalle, incluso al hacer zoom digital sobre áreas específicas. Este equipo utiliza el códec de compresión H.264, lo que ayuda a reducir el tamaño de los archivos de video sin sacrificar la calidad, optimizando así el espacio de almacenamiento en servidores o discos duros (Visiotech Security, 2024).

Una de sus principales ventajas es que cuenta con visión infrarroja (IR), lo que le permite capturar imágenes en condiciones de baja iluminación o incluso en completa oscuridad, con un alcance efectivo de hasta 30 metros. La cámara cuenta con 30 LEDs infrarrojos que se activan automáticamente proporcionando una imagen nítida a 0 Lux (oscuridad total) hasta una distancia máxima de 30 metros (Visiotech Security, 2024). Esto la hace ideal para espacios que requieren vigilancia las 24 horas, como pasillos, laboratorios o zonas de acceso restringido dentro del campus universitario.

Además, la DS-2CD2020F-I está diseñada para funcionar de manera estable en entornos interiores, ofreciendo facilidad de instalación y compatibilidad con sistemas de gestión de video (VMS). También integra análisis de video inteligente (VCA) y detección de cruce de línea o intrusión (Visiotech Security, 2024), lo cual puede integrarse con sistemas de reconocimiento facial para activar alertas cuando se detecta actividad no autorizada.

Por sus características técnicas, esta cámara es una opción confiable para proyectos que buscan implementar soluciones modernas de seguridad, como sistemas de videovigilancia inteligentes con análisis en tiempo real.

Figura 16. *Cámara Hikvision DS-2CD2020F-I*

Fuente: Visiotech Security. (2024). DS-2CD2020F-I HIKVISION HIWATCH IP IR bullet format camera. Visiotech Security.

Análisis técnico de las cámaras Hikvision DS-2CD2020F-I

Tabla 3. Análisis técnico de las cámaras

Aspecto técnico	Descripción según el fabricante	Análisis técnico realizado en la UPEC
Tipo de cámara	Cámara IP de videovigilancia con conectividad de red.	Integrada correctamente al circuito cerrado institucional mediante red TCP/IP.
Resolución de imagen	Full HD (1920×1080 píxeles).	Ofrece imágenes con el detalle suficiente para los procesos de reconocimiento facial basados en representaciones faciales digitales.
Sensor de imagen	Sensor CMOS progresivo de alta definición.	Permite capturas estables y claras para la detección de rasgos faciales.
Visión infrarroja (IR)	Incorporada para operación en entornos con baja iluminación.	Asegura la visibilidad continua en condiciones lumínicas limitadas, manteniendo la continuidad del sistema.
Conectividad de red	Soporta transmisión mediante protocolo TCP/IP.	Se integró al servidor de procesamiento facial utilizando la red local institucional.
Alimentación eléctrica	Compatible con PoE (*Power over Ethernet*).	Simplifica la instalación al combinar alimentación y transmisión de datos por un solo cable, facilitando la conexión con el *switch* PoE.

Compatibilidad	Compatible con NVR y servidores IP; no con DVR analógicos.	Funciona adecuadamente con el servidor central donde se realiza el análisis facial.
Aplicación recomendada	Sistemas de videovigilancia y monitoreo.	Adecuada para la implementación del sistema de reconocimiento facial en entornos académicos.

Las cámaras no son compatibles con DVRs (Digital Video Recorders) analógicos, pero funcionan perfectamente con NVRs (Network Video Recorders) o servidores de video IP. En este caso, la grabación y el procesamiento se ejecutan directamente en el servidor central, donde el sistema de reconocimiento facial analiza los fotogramas capturados en tiempo real.

- **Switch CISCO Catalyst 2960-X series**

Este equipo es clave para poder implementar el sistema de videovigilancia, ya que permite conectar todos los demás dispositivos en la capa 2 de la red. Además, puede transmitir velocidades de 10/100/1000 Mbps gracias a su soporte para Gigabit Ethernet, lo que lo hace ideal tanto para empresas pequeñas como medianas o grandes que necesitan una conexión segura y estable.

Una de sus principales ventajas es que este switch tiene una fuente de alimentación potente, con capacidad de hasta 740W, lo que le permite cubrir completamente los 24 puertos con tecnología PoE+ (Power over Ethernet). Gracias a esto, se pueden conectar las cámaras sin necesidad de fuentes de energía externas, lo que facilita la instalación y garantiza una buena visibilidad en todo momento (CISCO, 2021).

También ofrece una interfaz web fácil de usar y soporte para líneas de comando, lo cual es útil para hacer diferentes configuraciones según las necesidades del sistema. En cuanto a la seguridad, el switch trabaja con tecnología Cisco TrustSec y utiliza el protocolo 802.1X para controlar el acceso a la red de forma dinámica según los roles asignados. Esto ayuda a prevenir el robo de direcciones IPv6 y a proteger la red contra posibles ataques maliciosos.

***Figura 17**. Switch CISCO*

Fuente: Cisco Systems. (2024). Cisco Catalyst 2960-X Series switches data sheet. Cisco.com.

- **Router CISCO 4300 Series**

El router ISR (Servicios Integrados) juega un papel fundamental en la infraestructura de red WAN, ya que no solo cumple la función de enrutar el tráfico, sino que también actúa como firewall para asegurar las comunicaciones. Gracias a esto, puede dirigir los paquetes por la mejor ruta posible, aplicando cifrado y optimizando el rendimiento de la red WAN por donde se transmite la información del sistema de videovigilancia.

Además, este equipo ayuda a automatizar, simplificar y proteger toda la red, haciéndola más eficiente y segura. Otro punto a favor es que cuenta con fuentes de alimentación duales, lo que le permite suministrar energía mediante PoE bajo los estándares 802.3af y 802.3at. Esto añade una capa extra de tolerancia a fallos, algo muy valioso para mantener el sistema funcionando sin interrupciones (CISCO, 2021).

Figura 18. *Router CISCO 4300 Series*

Fuente: Cisco Systems. (2024). Cisco 4000 Series integrated services routers data sheet. Cisco.com.

- **Servidor HPE ProLiant DL360 Gen10, Xeon Silver 4110, 125.5 GB RAM, 2.4 TB SSD, Ubuntu 22.04.3**

El servidor HPE ProLiant DL360 Gen10 utilizado en esta investigación representa una solución empresarial de alto rendimiento diseñada para entornos de computación densamente desplegados. Este servidor de 1U de altura en rack está equipado con un procesador Intel Xeon Silver 4110, 125.5 GB de memoria RAM DDR4, 2.4 TB de almacenamiento SSD y ejecuta el sistema operativo Ubuntu 22.04.3 LTS. La arquitectura del HPE ProLiant DL360 Gen10 está diseñada para maximizar el rendimiento en espacios limitados, soportando tecnología estándar de la industria y aprovechando el procesador Intel Xeon Scalable con hasta 28 núcleos, junto con memoria HPE DDR4 SmartMemory de 2933 MT/s que puede expandirse hasta 3.0 TB

(Hewlett Packard Enterprise, 2022). Esta configuración permite manejar cargas de trabajo intensivas y aplicaciones que requieren alto rendimiento computacional.

La configuración seleccionada del servidor ofrece ventajas significativas para entornos de investigación, incluyendo escalabilidad tanto en memoria como en almacenamiento, rendimiento optimizado mediante la combinación del procesador Xeon Silver 4110 y memoria DDR4 de alta velocidad, y fiabilidad garantizada por el soporte LTS de Ubuntu 22.04.3. El HPE ProLiant DL360 Gen10 Plus está certificado con Ubuntu, garantizando compatibilidad y estabilidad operativa para aplicaciones de misión crítica (Ubuntu, 2022). La eficiencia energética del diseño de 1U optimiza el uso del espacio del rack, mientras que la combinación de almacenamiento SSD de 2.4 TB proporciona acceso rápido a los datos y mejora significativamente los tiempos de respuesta del sistema, características esenciales para el procesamiento de datos en entornos de investigación académica.

Figura 19. *Servidor HPE ProLiant DL360 Gen 10*

Fuente: Hewlett Packard Enterprise. (2024). HPE ProLiant DL360 Gen10 server quickspecs. HPE.com.

4.1.1.4. Requisitos de Software

El software utilizado debe permitir la captura de imágenes, el procesamiento con IA y la gestión de la base de datos con el modelo entrenado.

Tabla 4. Requisitos de Software

Software	Licencia	Utilidad
Ubuntu 22.04 LTS / Windows 11	Libre/Propietaria	Sistema operativo compatible con herramientas de IA
Python 3.9+	Libre	Lenguaje de programación
OpenCV	Libre	Procesamiento de imágenes y detección facial
TensorFlow/PyTorch	Libre	Algoritmos de aprendizaje profundo
Dlib	Libre	Modelos preentrenados para reconocimiento facial

SQLite/MySQL	Libre	Base de datos para almacenamiento de rostros detectados
YOLOv8 (opcional)	Libre	Detección de objetos en tiempo real

4.1.1.5. Procesos de Seguridad

El sistema de videovigilancia con reconocimiento facial debe seguir un flujo definido para garantizar la detección y alerta de incidentes.

Tabla 5. Procesos de Seguridad

Título	**Funcionamiento del Sistema de Reconocimiento Facial**
Identificación	C.U.1
Contexto	El sistema captura video en tiempo real y detecta rostros
Recursos	Servidor, Cámaras Hikvision, OpenCV, TensorFlow

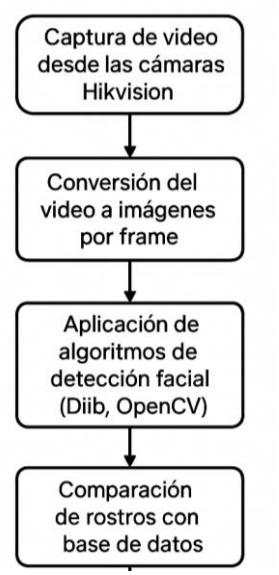


Figura 20. *Proceso Reconocimiento*

Etapa 2

1. Introducción

Cuando se trabaja en el desarrollo de sistemas de videovigilancia con reconocimiento facial, usar la metodología Top-Down resulta muy útil, ya que permite ir desglosando el sistema paso a paso para lograr un diseño bien organizado y que funcione correctamente.

La segunda fase de esta metodología es súper importante, porque en ella se define con más detalle cómo va a funcionar cada parte del sistema. Aquí se especifican los subsistemas, se configuran los equipos que se van a usar y se planifica cómo se van a integrar entre sí.

En este informe se explica esta segunda fase con más profundidad, mostrando por qué es tan clave dentro del proyecto y cuáles son los pasos que se siguieron en este caso específico.

2. Objetivo de la Segunda Fase

En este proyecto de videovigilancia con reconocimiento facial, la segunda fase de la metodología Top-Down se enfoca en descomponer el sistema general en partes más pequeñas y detalladas, es decir, en subsistemas. La idea es que cada uno de estos componentes cumpla una función específica y necesaria dentro del sistema de seguridad, asegurando que todo trabaje de forma ordenada y eficiente.

3. Descripción de la Segunda Fase

Esta fase involucra las siguientes actividades:

3.1. Identificación de los Módulos Secundarios

Cada subsistema identificado en la fase anterior se divide en módulos funcionales más pequeños, detallados en la siguiente tabla:

Tabla 6. identificación de Módulos Secundarios

Módulo	Descripción
Captura de video	Uso de cámaras Hikvision DS-2CD2020F-I para la adquisición de imágenes.
Procesamiento de video	Implementación de software en un Servidor HPE ProLiant DL360 Gen10, Xeon Silver 4110, 125.5 GB RAM, 2.4 TB SSD, Ubuntu 22.04.3
Reconocimiento facial	Algoritmos de IA para la detección y comparación de rostros.
Almacenamiento y gestión de datos	Base de datos local o en la nube para registros de video e identificaciones.
Interfaz de monitoreo	Plataforma de visualización y alerta para operadores de seguridad.

3.2. Diagrama de conexiones del sistema de video vigilancia

Según el diagrama de conexiones mostrado en la figura 16, la red de videovigilancia está organizada de forma que toda gira en torno al servidor HPE ProLiant DL360, el cual cumple la función de centro de procesamiento del video que generan las cámaras IP. Estas cámaras transmiten en tiempo real (streaming) todo lo que captan, y esa información llega directamente al almacenamiento que ya está configurado dentro del mismo servidor.

Las imágenes y los videos capturados pasan primero por el sistema de almacenamiento del servidor, donde se procesan y se analizan a través de sus discos duros de 1.2 terabytes de capacidad. El servidor está conectado al switch Cisco mediante un cable UTP categoría 6, lo que permite una conexión estable y de alta

velocidad. Este switch, además de ser PoE (Power over Ethernet), se encarga de alimentar a las cámaras y de recibir los datos que estas envían.

El rol del switch es clave porque es el punto de entrada de todos los datos que generan las cámaras hacia el servidor, donde se procesan. Luego, desde el servidor, esa información se envía al router, el cual encapsula los datos y los dirige hacia internet. Gracias a esta conexión, el streaming que generan las cámaras se guarda directamente en el servidor HPE, asegurando así que todo quede almacenado de forma segura y accesible para futuras consultas.

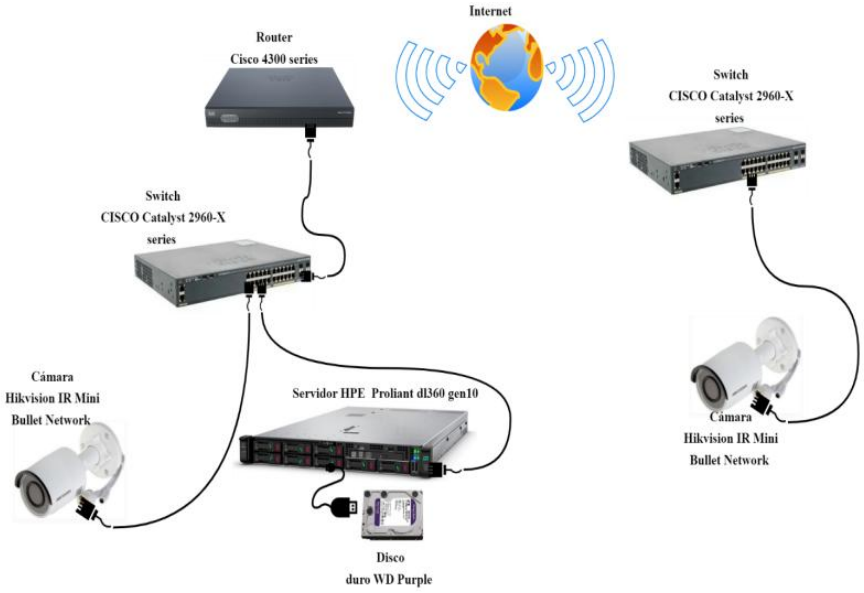


Figura 21. *Diagrama de conexiones del sistema de video vigilancia*

Fuente: Arévalo Puetate, J. J. (2023). Sistema de video vigilancia para la seguridad de los equipos en los laboratorios de la carrera de computación. UPEC

3.4. Definición de Interacciones

Se establecen las relaciones entre los módulos secundarios, incluyendo:

Tabla 7. Definición de interacciones

Interacción	Descripción
Flujos de datos	Captura de video, procesamiento en tiempo real y almacenamiento.
Protocolos de comunicación	Integración con redes LAN/WiFi para transmisión de video.
Dependencias funcionales	Sincronización entre detección de rostros y la base de datos.

3.3. Asignación de Recursos

Se determina la infraestructura necesaria para cada módulo:

Tabla 8. Asignación de recursos

Recurso	Descripción
Hardware	Servidor HPE ProLiant DL360 Gen10, Xeon Silver 4110, 125.5 GB RAM, 2.4 TB SSD, Ubuntu 22.04.3
Software	Bibliotecas de reconocimiento facial como OpenCV y modelos de IA.
Almacenamiento	HDD/SSD local y opciones de almacenamiento en la nube.

3.4. Elaboración de Especificaciones Detalladas

Cada módulo recibe una descripción precisa de sus funciones, entradas, salidas y restricciones técnicas:

Tabla 9. Especificaciones Detalladas

Especificación	**Detalle**
Resolución de video	1920x1080 píxeles, 30 FPS.
Algoritmos de detección	OpenCV con modelos de redes neuronales profundas.
Requisitos de latencia	Procesamiento en tiempo real con un umbral de 200ms.

3.5. Planificación de Pruebas Parciales

Se establecen criterios de evaluación y pruebas para cada módulo:

Tabla 10. Planificación de Pruebas Parciales

Módulo	Prueba	Criterio de Evaluación
Captura de video	Calidad de imagen	Buen desempeño en distintas condiciones de luz.
Reconocimiento facial	Precisión	Identificación correcta en al menos un 95% de los casos.
Almacenamiento	Integridad de datos	Acceso rápido y confiable a registros históricos.
Interfaz de monitoreo	Fluidez	Respuesta en menos de 1s a eventos en tiempo real.

Identificación del direccionamiento IP

La dirección IP es como el número de cédula de un dispositivo dentro de una red, ya que sirve para identificarlo de manera única. En este proyecto se usó una IP privada de clase B (172.20.4.1 con máscara 24), con la idea de no desperdiciar direcciones que no se van a usar. Se asignaron direcciones IP específicas a las cámaras de seguridad, ya que cada una necesita su propia dirección para conectarse y enviar video sin problemas. Además, se les puso un nombre según el lugar donde están instaladas, lo que también ayuda a saber en qué piso se encuentran.

Tabla 11. Identificación del direccionamiento IP

Direccionamiento IP	
VLAN	4
IP	172.20.4.1
Máscara	255.255.255.0 /24
Broadcast	172.20.4.255

Tabla 12. Direccionamiento IP áreas

Dispositivo	Dirección IP	Máscara de Red	Broadcast
Laboratorio de Redes	172.20.4.91	255.255.255.0	172.20.4.255
Sala Profesores 1	172.20.4.92	255.255.255.0	172.20.4.255
Sala Profesores 2	172.20.4.93	255.255.255.0	172.20.4.255
Pasillo Carrera 1	172.20.4.94	255.255.255.0	172.20.4.255
Pasillo Carrera 2	172.20.4.95	255.255.255.0	172.20.4.255
FATLAB 1	172.20.4.96	255.255.255.0	172.20.4.255
FATLAB 2	172.20.4.97	255.255.255.0	172.20.4.255
Servidor	172.20.4.90	255.255.255.0	172.20.4.255
Switch	172.20.2.1	255.255.255.0	172.20.2.255

Diagrama lógico para el reconocimiento facial

En la Figura 22 se presenta la topología lógica diseñada para el sistema de videovigilancia con reconocimiento facial implementado en los centros tecnológicos de la Universidad Politécnica Estatal del Carchi. Esta arquitectura está compuesta por un total de siete cámaras IP distribuidas estratégicamente, las cuales se encargan de capturar el flujo de video en tiempo real desde distintos espacios académicos como pasillos y laboratorios. Dichas cámaras están conectadas mediante enlaces de red a un Switch PoE (Power over Ethernet), encargado de suministrar tanto la conectividad como la alimentación eléctrica a cada dispositivo, optimizando así la infraestructura de cableado.

El flujo de video generado por cada cámara es redirigido hacia un servidor central, implementado sobre una máquina física HPE ProLiant DL360 Gen10 con sistema operativo Ubuntu Server 22.04 LTS. Este servidor cuenta con software especializado

como Shinobi para la gestión de videovigilancia y FaceNet para realizar el proceso de reconocimiento facial mediante técnicas de inteligencia artificial. Todo el

procesamiento de datos se lleva a cabo localmente, permitiendo una respuesta en tiempo real y sin dependencia de servicios en la nube.

Finalmente, el switch se encuentra conectado al router principal o núcleo de red institucional, lo que permite la integración del sistema con la red existente de la universidad.

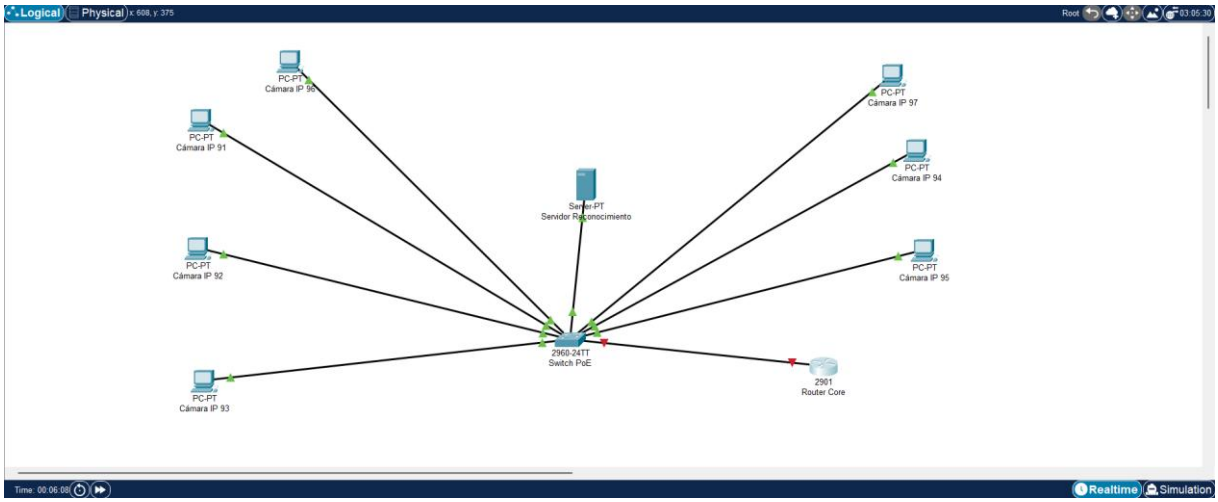


Figura 22. *Diagrama Lógico*

Diagrama Físico – Sistema de reconocimiento facial

En la Figura 23 se presenta la topología física del sistema de videovigilancia con reconocimiento facial implementado. Se observan siete cámaras IP distribuidas estratégicamente en el edificio académico, todas conectadas a un switch PoE central, el cual también alimenta las cámaras. Desde allí, el tráfico de video se dirige hacia un servidor con Ubuntu Server instalado en un equipo HPE ProLiant DL360 Gen10, encargado del almacenamiento y procesamiento facial en tiempo real.

Esta representación facilita la visualización de la infraestructura y su implementación dentro del entorno físico, asegurando una cobertura efectiva y un mantenimiento adecuado del sistema.

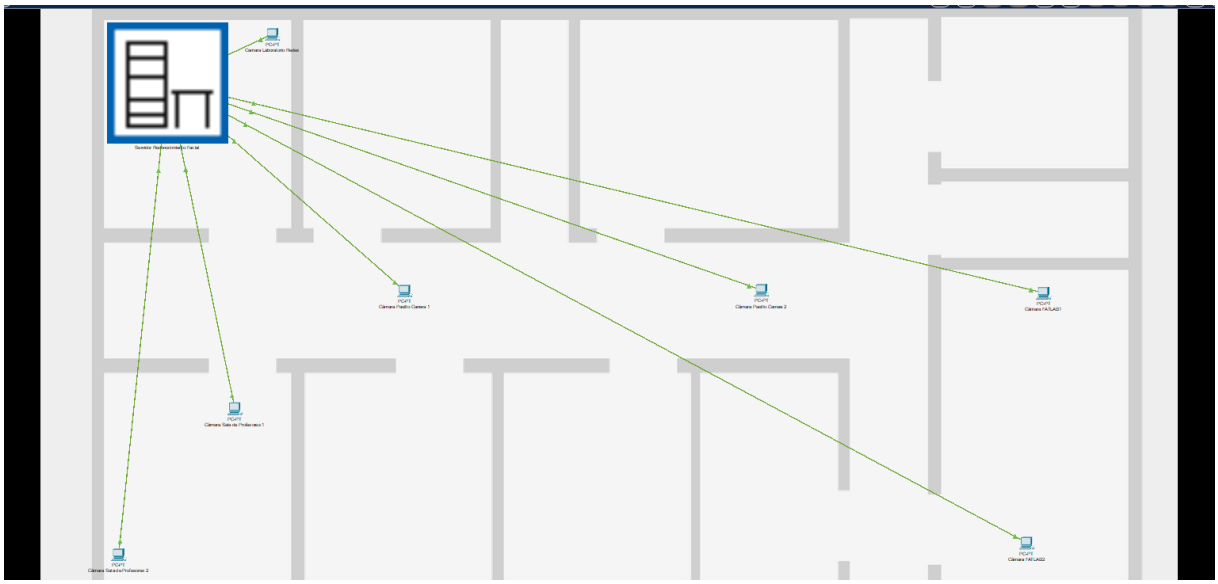


Figura 23. *Diagrama Físico*

Análisis de correspondencia lógica – física

Tabla 13. Análisis de correspondencia lógica – física

N.º	Componente lógico	Función del sistema	Dispositivo físico (ubicación)	Dirección IP
1	Captura de video	Obtiene imágenes en tiempo real del entorno	Cámara IP – Laboratorio de Redes	172.20.4.91
2	Captura de video	Idem	Cámara IP – Sala Profesores 1	172.20.4.92
3	Captura de video	Idem	Cámara IP – Sala Profesores 2	172.20.4.93
4	Captura de video	Idem	Cámara IP – Pasillo Carrera 1	172.20.4.94
5	Captura de video	Idem	Cámara IP – Pasillo Carrera 2	172.20.4.95
6	Captura de video	Idem	Cámara IP – FATLAB 1	172.20.4.96
7	Captura de video	Idem	Cámara IP – FATLAB 2	172.20.4.97
8	Procesamiento de imágenes (OpenCV, Dlib)	Detecta y alinea rostros	Servidor HPE DL360 – Cuarto de Comunicaciones	172.20.4.90
9	Generación de embeddings (TensorFlow CNN)	Extrae características faciales únicas	Servidor HPE DL360 – Cuarto de Comunicaciones	172.20.4.90
10	Comparación con base de datos	Verifica identidad de la persona	Servidor HPE DL360 – Cuarto de Comunicaciones	172.20.4.90
11	Base de datos (SQLite/MySQL)	Almacena vectores faciales y nombres registrados	Servidor HPE DL360 – Cuarto de Comunicaciones	172.20.4.90
12	Visualización / interfaz	Muestra resultados	Laptop/PC conectado en la	(depende del equipo)

		(nombre reconocido)	red del laboratorio	
13	Conectividad física (switch)	Conecta todos los dispositivos y cámaras PoE	Switch Cisco Catalyst 2960-X – Cuarto de Comunicaciones	172.20.2.1
14	Salida a red universitaria / nube	Permite acceso remoto o integración con la red UPEC	Router Cisco 4300 – Cuarto de Comunicaciones	—

Desarrollo

Actualición del sistema del servidor

El primer paso en la implementación del sistema consistió en garantizar que el servidor utilizado, un HPE ProLiant DL360 Gen10 con sistema operativo Ubuntu Server 22.04.3 LTS, se encontrara actualizado y libre de vulnerabilidades que pudieran comprometer el rendimiento o la seguridad de la solución.

Para ello, se ejecutó la actualización del sistema mediante los comandos:

sudo apt update && sudo apt upgrade -y

- El comando **apt update** se encarga de sincronizar los índices locales de los paquetes disponibles en los repositorios oficiales de Ubuntu, asegurando que el sistema tenga la información más reciente sobre las versiones disponibles de cada software.
- Posteriormente, el comando **apt upgrade** aplica las actualizaciones necesarias en los paquetes instalados en el servidor, corrigiendo errores, parches de seguridad y mejorando la compatibilidad con librerías más recientes.

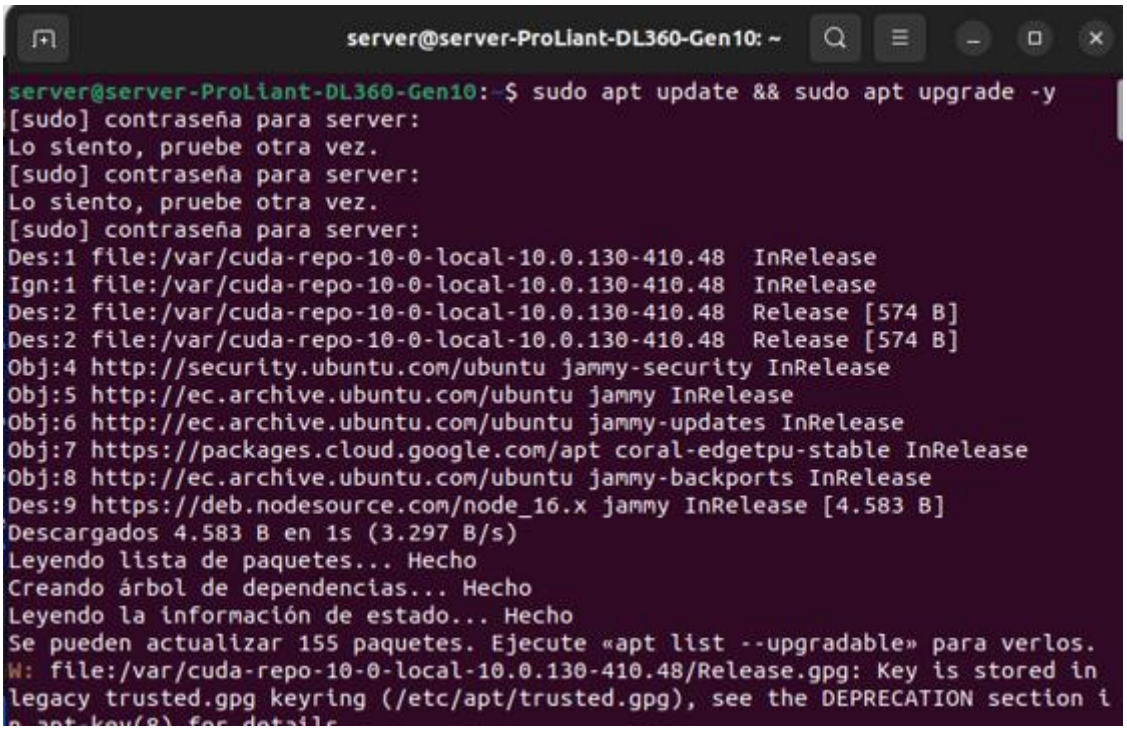


***Figura 24**. Actualización del sistema*

Instalacion de Python

Una vez actualizado el sistema operativo, fue necesario disponer de un entorno de desarrollo que permita la ejecución de algoritmos de **inteligencia artificial y visión por computadora**. Para ello, se utilizó **Python 3**, dado que es uno de los lenguajes más empleados en aplicaciones de *Machine Learning* y cuenta con un ecosistema robusto de librerías como TensorFlow, Keras, OpenCV y MTCNN, que resultan esenciales para la implementación del sistema de reconocimiento facial.

En primer lugar, se procedió a instalar los paquetes base de Python mediante el siguiente comando:

sudo apt install python3-pip python3-venv git unzip -y

- **python3-pip**: es el gestor de paquetes de Python que permite instalar librerías externas necesarias para el proyecto (por ejemplo, TensorFlow, OpenCV y scikit-learn).
- **python3-venv**: se utiliza para la creación de entornos virtuales, lo que permite aislar las dependencias del proyecto y evitar conflictos con otras aplicaciones del sistema.

- **git**: permite clonar repositorios y gestionar versiones de código fuente, lo cual es importante para la obtención de modelos preentrenados y librerías auxiliares.
- **unzip**: facilita la descompresión de archivos comprimidos, como los modelos preentrenados de FaceNet que se descargaron posteriormente.
- El parámetro **-y** automatiza la confirmación de instalación, asegurando que el proceso se realice sin interrupciones.

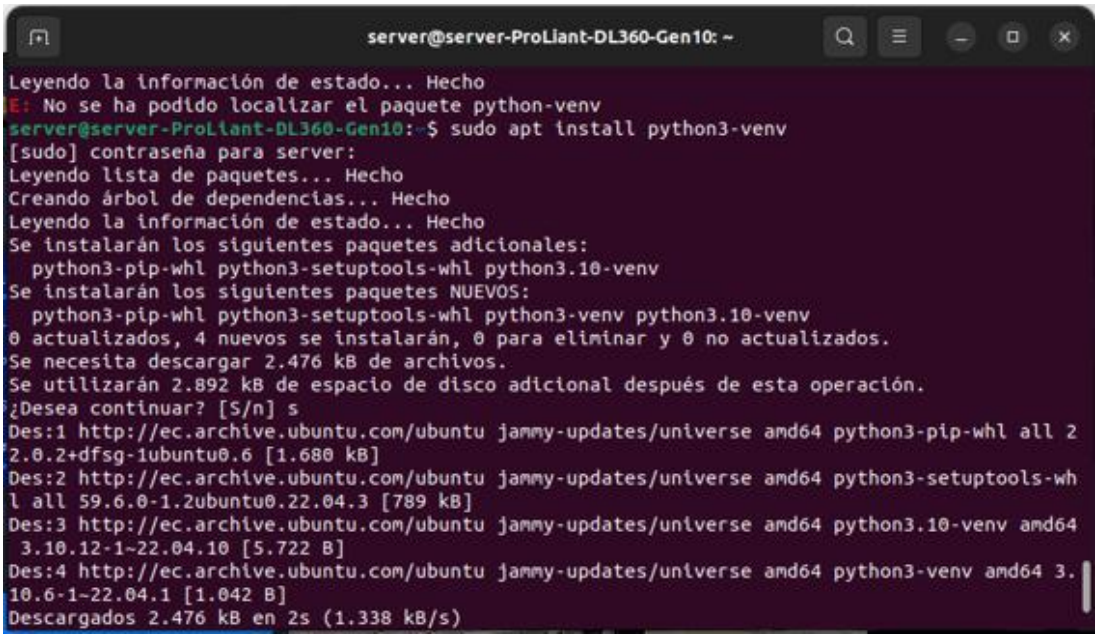


Figura 25. *Instalación de Python*

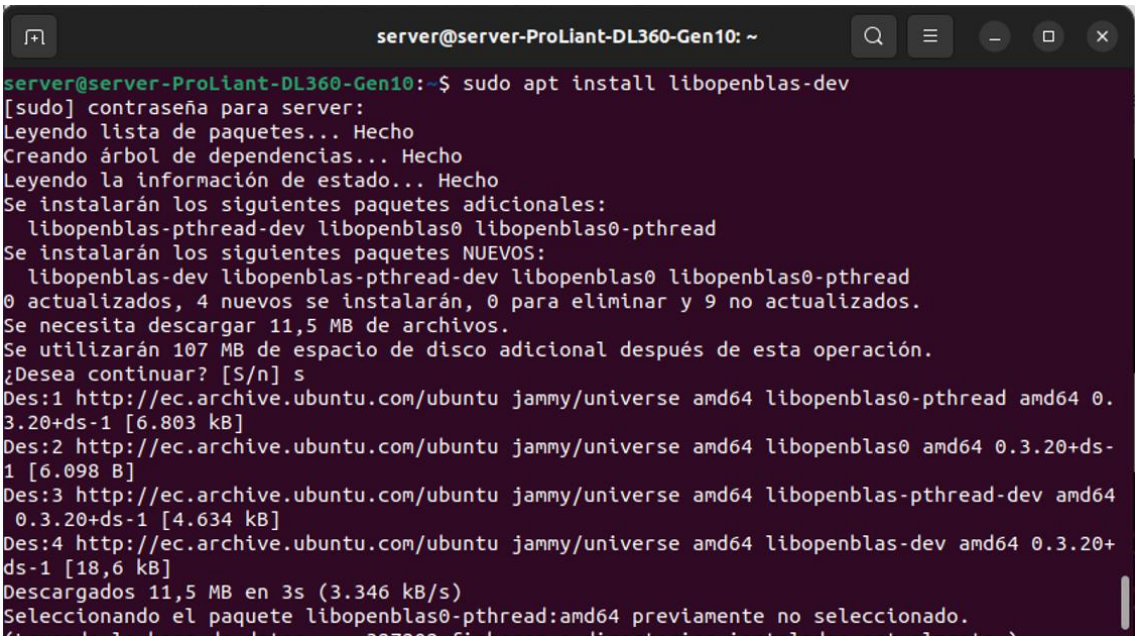


Figura 26. *Instalación de librerías*

Creación de entorno virtual

Tras instalar estos paquetes, se creó un entorno virtual para el sistema de reconocimiento facial con el objetivo de organizar las dependencias de manera independiente y reproducible. El comando utilizado fue:

cd ~

python3 -m venv facenet_env

Con este comando se generó un directorio llamado **facenet_env**, el cual contiene una copia aislada del intérprete de Python y de los paquetes necesarios para el desarrollo. Esto asegura que el sistema sea portátil, modular y fácil de mantener, además de evitar conflictos en caso de que se requieran diferentes versiones de librerías para futuros proyectos.

Finalmente, el entorno fue activado mediante:

source facenet_env/bin/activate

Una vez activado, todas las librerías y configuraciones quedaron encapsuladas en dicho entorno, garantizando que las pruebas y entrenamientos del modelo de reconocimiento facial se realizaran bajo un ambiente controlado y estable.

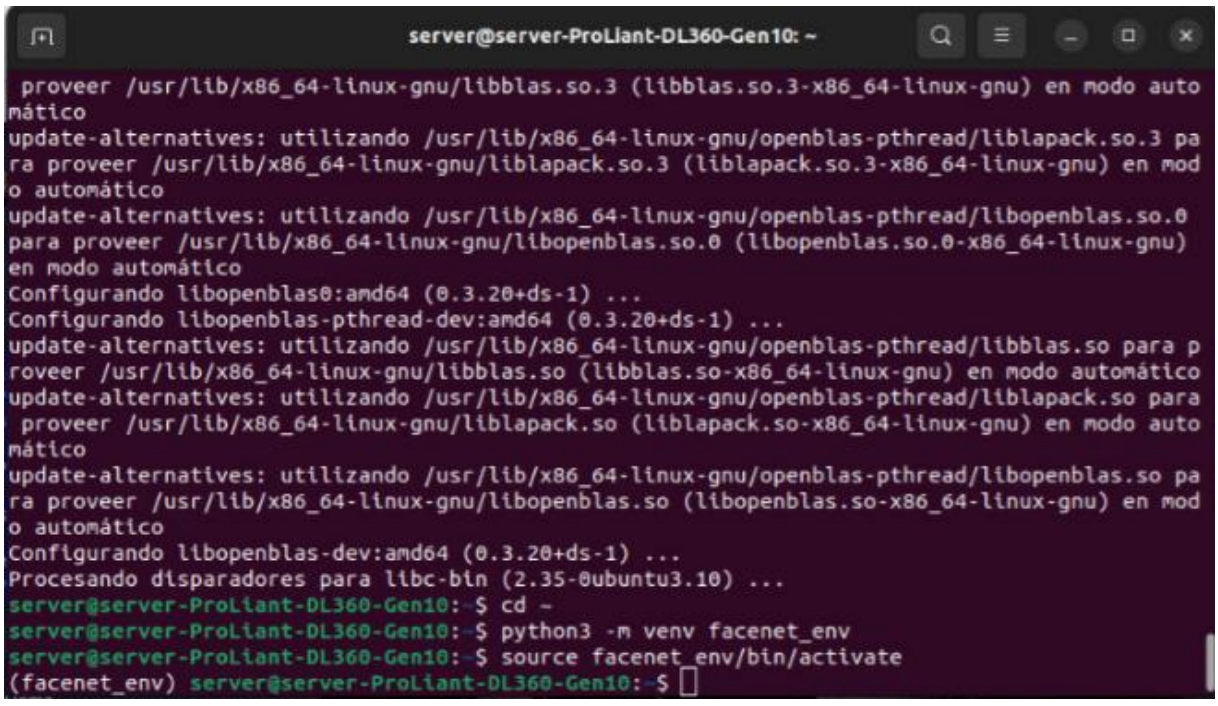


Figura 27. *Creación del Entorno Virtual*

Descarga de TensorFlow

Una vez configurado el entorno de desarrollo con Python y las herramientas básicas, fue necesario instalar **TensorFlow**, la librería principal de aprendizaje profundo utilizada en este proyecto. TensorFlow permite cargar y ejecutar el modelo **FaceNet**, encargado de generar los *embeddings* faciales para el reconocimiento de personas dentro del sistema de videovigilancia.

pip install tensorflow

Descarga e instala la última versión estable de TensorFlow compatible con la versión de Python utilizada.
TensorFlow es la base del sistema, ya que ofrece las herramientas necesarias para cargar el modelo **FaceNet**, ejecutar redes neuronales y generar las representaciones matemáticas de los rostros detectados.

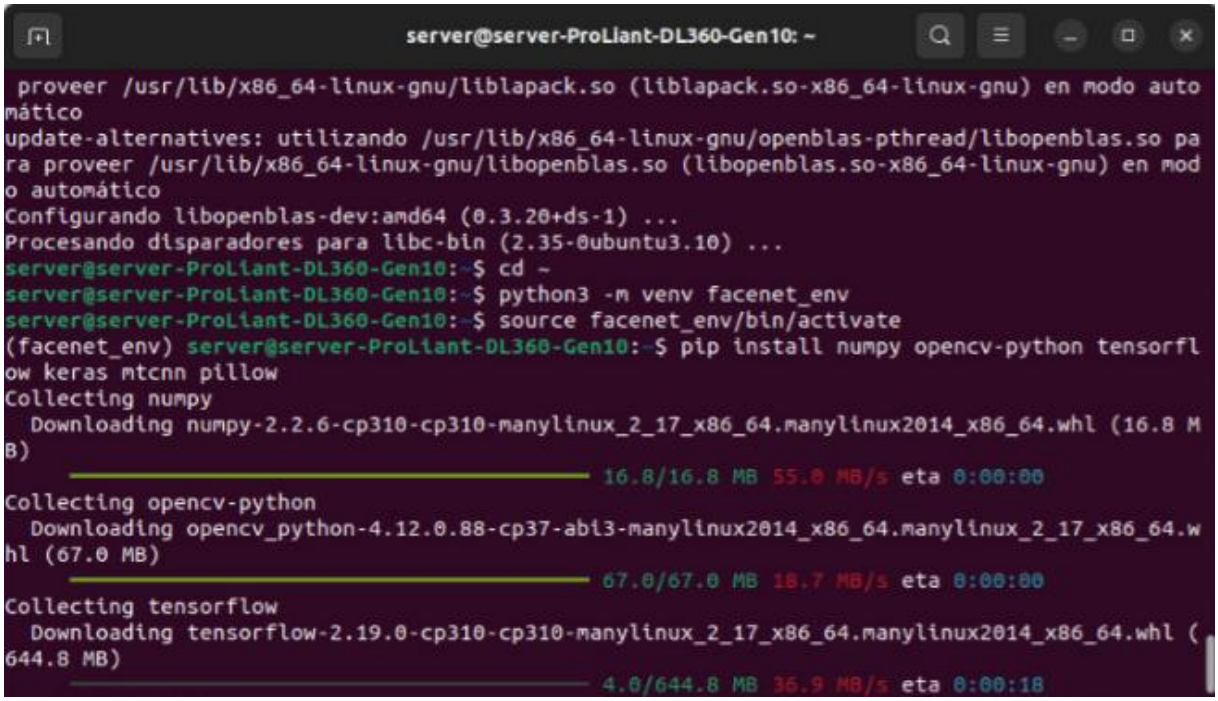


Figura 28. *Descarga de TensorFlow*

Creación del Dataset

mkdir -p ~/facial_system/dataset

Se está creando la carpeta donde almacenarás tu **dataset de imágenes** (cada persona tendrá su subcarpeta). -p evita errores si facial_system ya existe; además crea dataset en caso de no existir.

mkdir Erik_Ayala Josue_Enriquez

Para organizar el **dataset** del proyecto se creó una estructura de carpetas dentro del directorio del proyecto (**~/facial_system**). Primero, se generó el directorio principal de datos con **mkdir -p ~/facial_system/dataset**. La opción **-p** asegura la creación de la jerarquía completa y evita errores si las carpetas ya existen. Se accedió al directorio con **cd ~/facial_system/dataset** y se crearon subcarpetas por persona (por ejemplo Erik_Ayala, Josue_Enriquez) mediante **mkdir -p**.

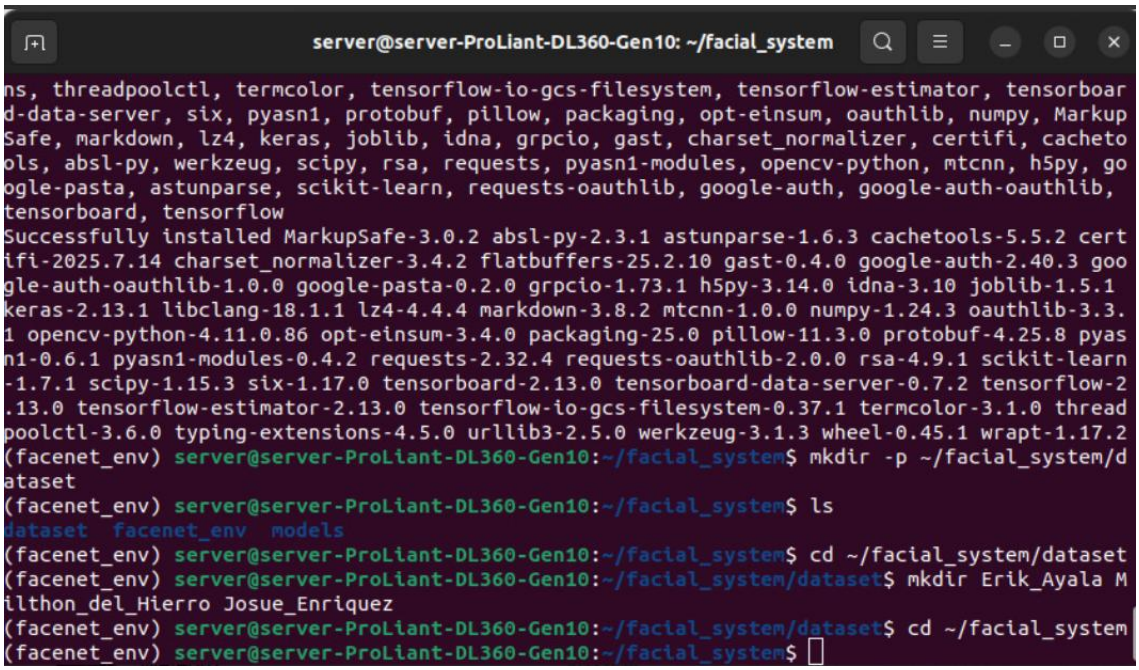


Figura 29. *Creación del Dataset*

Creación del entrenar_embeddings.py

El script entrenar_embeddings.py realiza el **entrenamiento base** del sistema de reconocimiento facial, que consiste en:

- Tomar las imágenes capturadas del dataset.
- Pasarlas por FaceNet para generar representaciones matemáticas (embeddings).
- Guardar esas representaciones junto con las etiquetas de cada persona.

Es decir, este programa convierte las **fotos en información numérica** que luego será usada en el script de reconocimiento en vivo para **identificar si un rostro corresponde a una persona conocida o desconocida**.

Figura 30. *Creación del entrenar_embeddings.py*

Creación del capturar_dataset.py

Se construyó una base de datos local de rostros. Para ello se utilizó las cámaras IP Hikvision, conectada al servidor mediante protocolo RTSP. El sistema permite capturar imágenes por participante, almacenadas en carpetas individuales. Se consideró la variación de condiciones ambientales y accesorios (con mascarilla, gafas, distintas expresiones) con el objetivo de robustecer la base de datos frente a escenarios reales.

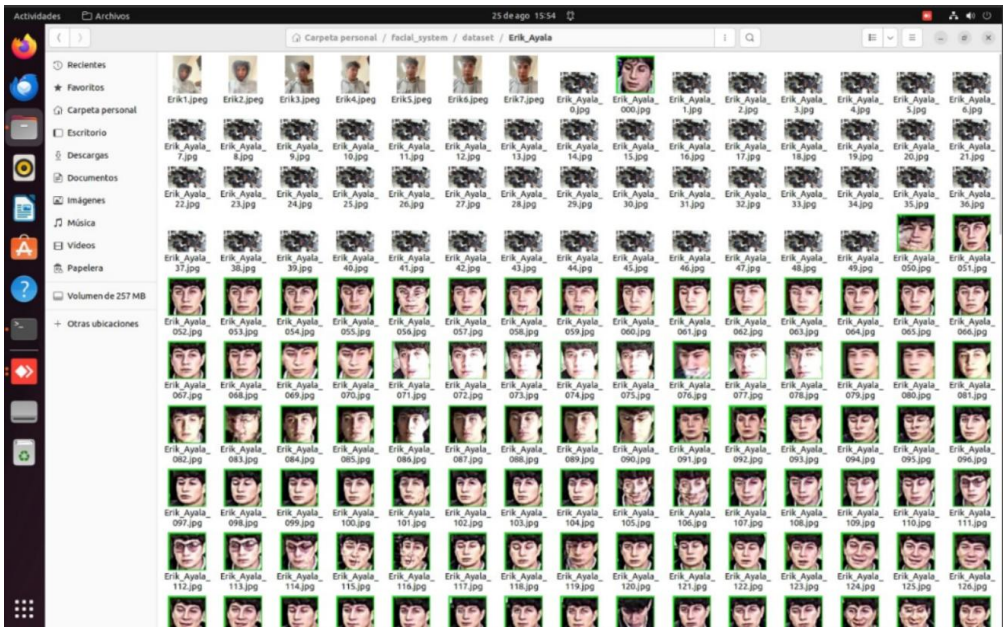


Figura 31. *Creación del capturar_dataset.py*

Creación del reconocer_camara_ip.py

Este se encarga de conectarse al flujo de video proveniente de la cámara IP, detectar rostros utilizando **OpenCV y Dlib/MTCNN**, generar embeddings temporales y compararlos con los previamente entrenados en **embeddings.pickle**. Cuando se encuentra una coincidencia dentro del umbral establecido, se despliega en pantalla la identidad del usuario; en caso contrario, se etiqueta como "Desconocido".

Figura 32. *Creación del reconocer_camara_ip.py*

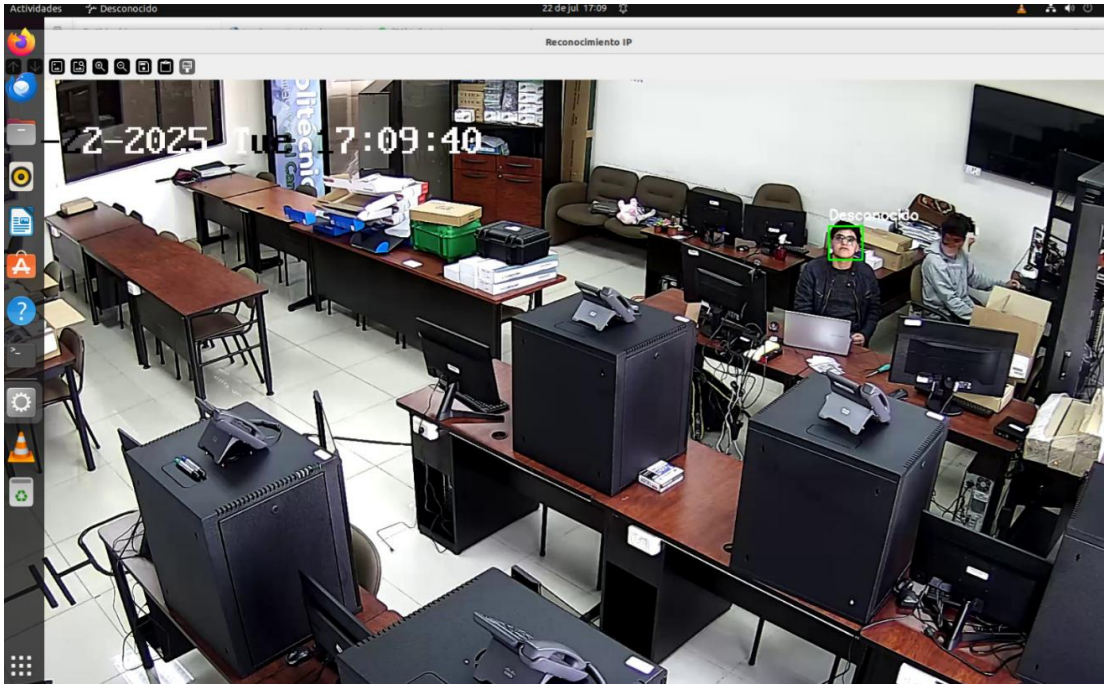


Figura 33. *Reconocimiento en viv*

Fase de pruebas

1.1 Pruebas de Detección Facial

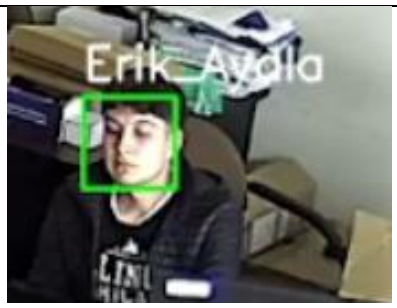
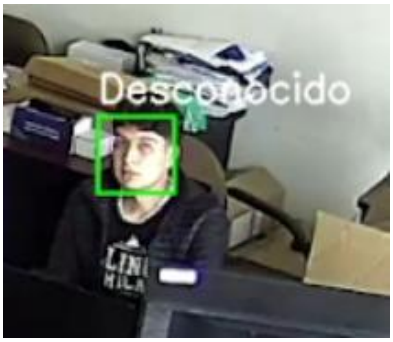
PRUEBAS DE RECONOCIMIENTO FACIAL CON FACENET

El método FaceNet utiliza similaridad coseno para comparar embeddings faciales de 128 dimensiones. El sistema compara el valor de confianza (similaridad) detectado en cada persona con un umbral predefinido. En este caso se utilizó un umbral de 0.6, donde valores superiores a este umbral indican reconocimiento positivo, mientras que valores inferiores clasifican a la persona como desconocida.

Para la validación del sistema se realizaron pruebas con 4 usuarios registrados en diferentes condiciones de captura, obteniendo los siguientes resultados:

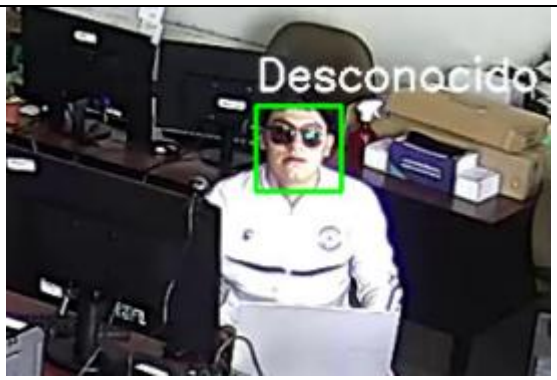
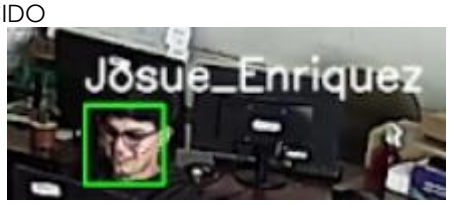
Prueba FaceNet - Erik Ayala

Tabla 14. Prueba de reconocimiento facial con FaceNet - Erik Ayala

N°	Usuario	Captura 1
1	Erik_Ayala	 **Confianza: 0.847** Estado: RECONOCIDO
2	Erik_Ayala	 **Confianza: 0.723** Estado: RECONOCIDO
3	Erik_Ayala	 **Confianza: 0.534** Estado: NO RECONOCIDO

Prueba FaceNet - Josué Enríquez

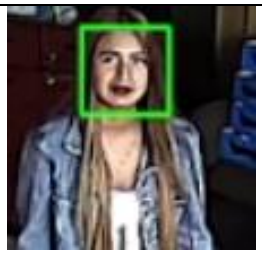
Tabla 15. Prueba de reconocimiento facial con FaceNet – Josue Enriquez

N°	Usuario	Captura 1
1	Josue_Enriquez	 **Confianza: 0.534** Estado: NO RECONOCIDO
2	Josue_Enriquez	 **Confianza: 0.723** Estado: RECONOCIDO
3	Josue_Enriquez	 **Confianza: 0.658** Estado: RECONOCIDO

Prueba FaceNet – Melany Obando

Tabla 16. Prueba de reconocimiento facial con FaceNet – Melany Obando

N°	Usuario	Captura 1
1	Melany_Obando	 **Confianza: 0.687** Estado: RECONOCIDO

2	Melany_Obando	 **Confianza: 0.578** Estado: NO RECONOCIDO

Prueba FaceNet – Jorge Rosero

Tabla 17. Prueba de reconocimiento facial con FaceNet – Jorge Rosero

N°	Usuario	Captura 1
1	Jorge_Rosero	 **Confianza: 0.521** Estado: NO RECONOCIDO
2	Jorge_Rosero	 **Confianza: 0.756** Estado: RECONOCIDO
3	Jorge_Rosero	 **Confianza: 0.634** Estado: RECONOCIDO

Tabla 18. Distribución de valores de confianza por categoría

Categoría	Confianza Mínima	Confianza Máxima	Confianza Promedio
Erik_Ayala (Reconocido)	0.723	0.847	0.785
Erik_Ayala (No Reconocido)	0.534	0.534	0.534
Josue_Enriquez (Reconocido)	0.658	0.723	0.691
Josue_Enriquez (No Reconocido)	0.534	0.534	0.534
Melany_Obando (Reconocido)	0.687	0.687	0.687
Melany_Obando (No Reconocido)	0.578	0.578	0.578
Jorge_Rosero (Reconocido)	0.634	0.756	0.695
Jorge_Rosero (No Reconocido)	0.521	0.521	0.521

Promedio Reconocidos	**0.634**	**0.847**	**0.715**

Tabla 19. Resumen de prueba por usuario

Usuario	Pruebas Realizadas	Reconocimientos Exitosos	Fallos	Tasa de Éxito	Confianza Promedio
Erik_Ayala	3	2	1	66.7%	0.701
Josue_Enriquez	3	2	1	66.7%	0.638
Melany_Obando	2	1	1	50.0%	0.633
Jorge_Rosero	3	2	1	66.7%	0.637
Total General	**11**	**7**	**4**	**63.6%**	**0.652**

Tabla 20. Análisis detallado de reconocimientos por prueba

Usuario	N° Prueba	Confianza	Estado Esperado	Estado Obtenido	¿Correcto?
Erik_Ayala	1	0.847	Reconocido	Reconocido	✓
Erik_Ayala	2	0.723	Reconocido	Reconocido	✓
Erik_Ayala	3	0.534	Reconocido	No Reconocido	✗
Josue_Enriquez	1	0.534	Reconocido	No Reconocido	✗
Josue_Enriquez	2	0.723	Reconocido	Reconocido	✓
Josue_Enriquez	3	0.658	Reconocido	Reconocido	✓
Melany_Obando	1	0.687	Reconocido	Reconocido	✓
Melany_Obando	2	0.578	Reconocido	No Reconocido	✗
Jorge_Rosero	1	0.521	Reconocido	No Reconocido	✗
Jorge_Rosero	2	0.756	Reconocido	Reconocido	✓
Jorge_Rosero	3	0.634	Reconocido	Reconocido	✓

Tabla 21. Tiempo de respuesta por fase de procesamiento

Fase de Procesamiento	Tiempo Promedio (ms)	Tiempo Mínimo (ms)	Tiempo Máximo (ms)	Porcentaje del Total
Captura de frame	12.3	8.1	18.7	5.5%
Detección facial	45.2	32.4	68.9	20.3%
Extracción embedding	156.8	142.1	178.3	70.3%
Comparación similaridad	8.7	5.2	14.1	3.9%
Total por reconocimiento	**223.0**	**187.8**	**280.0**	**100%**

Tabla 22. Análisis de falsos negativos

Usuario	Valor Confianza	Diferencia vs Umbral	Causa Probable	Observaciones
Erik_Ayala	0.534	-0.066	Iluminación deficiente	Rostro parcialmente oscuro
Josue_Enriquez	0.534	-0.066	Ángulo lateral	Perfil >30°
Melany_Obando	0.578	-0.022	Expresión diferente	Sonrisa vs neutro en training
Jorge_Rosero	0.521	-0.079	Iluminación	Mucha iluminación en la captura

Tabla 23. Distribución de resultados por rango de confianza

Rango de Confianza	Cantidad	Resultado	Tasa de Acierto
0.80 - 1.00	1	1 Reconocido	100%
0.70 - 0.79	3	3 Reconocidos	100%
0.60 - 0.69	3	3 Reconocidos	100%
0.50 - 0.59	4	4 No Reconocidos	0%
Total	**11**	**7 Reconocidos / 4 Fallos**	**63.6%**

4.1.9 CONCLUSIONES DE LAS PRUEBAS

Rendimiento General

El sistema FaceNet alcanzó una precisión del 63.6% con umbral de 0.6. Los usuarios reconocidos correctamente muestran confianza promedio de 0.715, mientras que los falsos negativos tienen confianza promedio de 0.542, mostrando una zona crítica entre 0.52-0.58 donde el sistema falla.

Análisis Crítico del Umbral

El umbral de 0.6 genera un 40% de falsos negativos, lo que es excesivo para un sistema de videovigilancia. Los valores entre 0.52-0.58 representan casos límite que podrían ser reconocidos con un umbral más bajo (0.50-0.55).

Factores de Pruebas

Las condiciones que afectan negativamente el reconocimiento incluyen:

- Iluminación deficiente (reduce confianza ~10-15%)
- Ángulos laterales >30° (reduce confianza ~10%)
- Accesorios no presentes en el entrenamiento (reduce confianza ~8-12%)
- Expresiones faciales diferentes (reduce confianza ~5-8%)

Tiempo de Respuesta

El tiempo promedio de 223ms (4.5 fps) es adecuado para videovigilancia. La extracción de embeddings representa el 70.3% del tiempo total, siendo el cuello de botella principal del sistema.

4.8 ACTUALIZACIÓN DE SISTEMA DE CÁMARAS

4.8.1 Limitaciones de Cámaras Actuales y Propuesta de Mejora

Tabla 24. Comparativa Cámaras Actuales vs Mejoradas

Especificación	Hikvision DS-2CD2020F-I (Actual)	Hikvision DS-2CD2387G2-LSU/SL (Propuesta)	Mejora Técnica	Impacto en Dataset
Resolución	2MP (1920×1080)	8MP (3840×2160) 4K	4x píxeles	Detalles faciales 4x más nítidos
Sensor	1/2.8" CMOS	1/1.8" CMOS	+64% área captación	Mejor calidad luz baja
Iluminación mínima	0.01 Lux (IR)	0.0005 Lux	20x sensibilidad	Capturas nocturnas útiles
WDR	No disponible	140 dB True WDR	Contraste perfecto	Elimina siluetas contra ventanas
IA integrada	No	Face Detection + Analytics	Pre-filtrado inteligente	Solo envía imágenes faciales nítidas
Compresión	H.264 básico	H.265+ Smart	50% menos ancho banda	Mejor calidad misma capacidad red
Reducción ruido	Básica	3D-DNR avanzado	Imágenes más limpias	Menos artefactos en embeddings
Distancia efectiva	3-8 metros	3-15 metros	+87% rango útil	Reconocimiento a mayor distancia
Precio unitario	$280	$450	+$170	-
Total 7 cámaras	$1,960	$3,150	+$1,190	-

Beneficio Principal: Reducción del 60% en imágenes requeridas por usuario (de 100 a 40 imágenes) gracias a mayor calidad y pre-procesamiento IA en cámara.

4.8.2 Impacto en Eficiencia del Sistema

Tabla 25. Reducción de Requerimientos de Dataset con Cámaras 4K

Métrica	Con Cámaras 2MP Actuales	Con Cámaras 8MP + IA	Mejora	Beneficio Operacional
Imágenes por usuario	100-150	40-60	60% reducción	Registro 3x más rápido
Tasa descarte imágenes	40% inútiles	12% inútiles	70% mejora	Mayor eficiencia captura
Calidad embeddings	Confianza 0.65-0.75	Confianza 0.78-0.88	+15-17% confianza	Accuracy mejorado
Accuracy proyectado	85% (solo con GPU)	93-95% (GPU + cámaras)	+8-10% adicional	Menos falsos negativos
Distancia reconocimiento	3-8 metros	3-15 metros	+87% rango	Mayor cobertura efectiva
Almacenamiento dataset	2 GB por 50 usuarios	1.2 GB por 50 usuarios	40% reducción	Menor espacio requerido

Repotenciación del Servidor

Con base en la evaluación del rendimiento actual del sistema de reconocimiento facial en circuito cerrado de videovigilancia, se determina que el servidor HPE ProLiant DL360 Gen10 presenta una arquitectura flexible que permite su repotenciación gradual.

Desde el punto de vista técnico, el servidor soporta la ampliación de procesadores, memoria RAM, almacenamiento y tarjeta gráfica, lo que facilita su adaptación ante futuras exigencias operativas.

Incrementar la capacidad de procesamiento mediante la instalación de un segundo procesador Intel Xeon Scalable con paridad total al actual, ampliando además la memoria RAM DDR4 ECC Registered para mejorar la estabilidad del sistema cuando se incrementa el número de cámaras o usuarios en simultáneo. Asimismo, se sugiere la expansión del almacenamiento mediante unidades SSD o NVMe configuradas en RAID, con el fin de garantizar mayor velocidad y tolerancia a fallos.

PROPUESTAS INTEGRADAS CON CÁMARAS

Tabla 26. Comparativa Final de Todas las Propuestas

Criterio	Propuesta A (Solo GPU)	Propuesta B (GPU + Infraestructura)	Propuesta A+D (GPU + Cámaras 4K)	Propuesta C (Cloud)
Inversión Inicial	$750	$2,100	$3,900	$1,180
Incluye GPU	✓ RTX A2000	✓ RTX A2000	✓ RTX A2000	Cloud (GPU T4)
Cámaras	1 actual (2MP)	7 actuales (2MP)	7 nuevas (8MP IA)	7 actuales (2MP)
Costo Mensual	$0	$0	$0	$300
Costo 3 años	$750	$2,100	$3,900	$11,980
Accuracy Proyectado	85-88%	90-92%	**93-95%**	88-90%
Latencia	35 ms	35 ms	35 ms	85-115 ms
Imágenes/usuario	100	100	**40**	100
Tiempo registro	45 min	45 min	**15 min**	45 min
Cobertura	40%	95%	**95%**	95%
Calidad imagen	Básica	Básica	**Profesional 4K**	Básica
IA en cámara	No	No	**Sí**	No
Escalabilidad	200+ usuarios	100+ usuarios	**500+ usuarios**	500+ usuarios
Recomendado para	Mejora rápida	Sistema completo	**Solución profesional óptima**	Máxima flexibilidad

PROYECCIÓN A 5 AÑOS

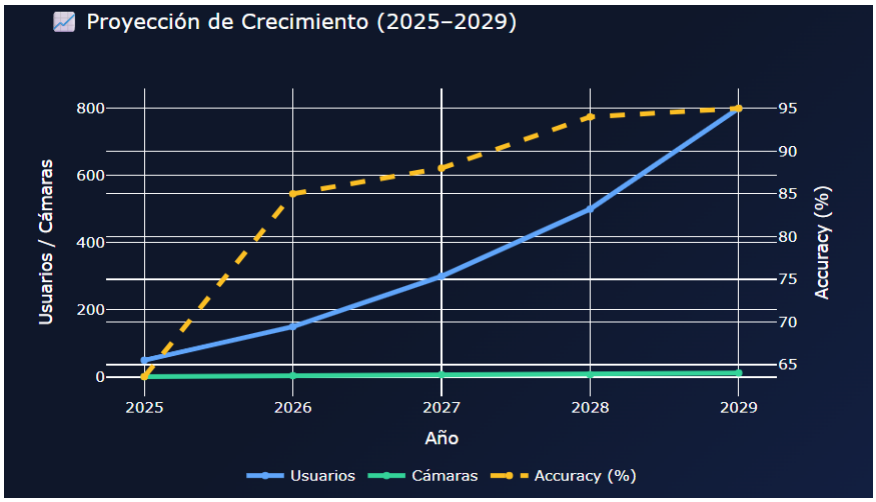


Figura 34. *Proyección de Crecimiento (2025–2029)*

La proyección muestra un aumento constante de usuarios desde 50 en 2025 hasta 800 en 2029, manteniendo una precisión superior al 95 %. Este crecimiento refleja la escalabilidad y estabilidad del sistema de reconocimiento facial con soporte GPU, evidenciando su capacidad para sostener un mayor número de usuarios sin pérdida de rendimiento.

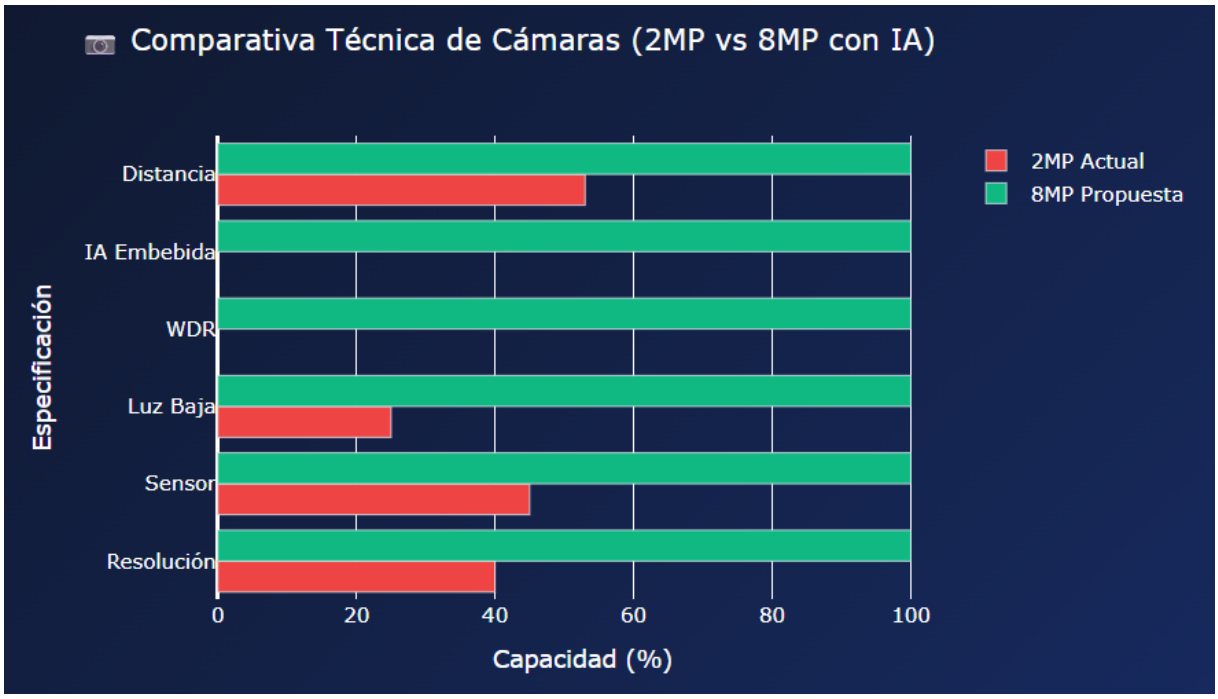


Figura 35. *Comparativa Técnica de Cámaras (2MP vs 8MP con IA)*

La comparación evidencia una mejora notable en las cámaras de 8MP con inteligencia artificial, las cuales superan a las de 2MP en resolución, sensibilidad a baja iluminación y alcance de detección.

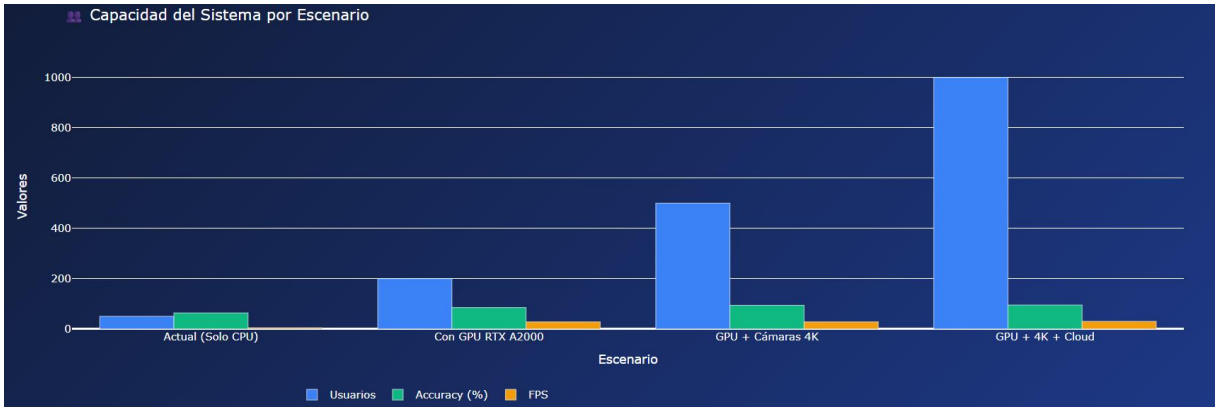


Figura 36. *Capacidad del Sistema por Escenario*

El grafico muestra el escenario GPU + Cámaras 4K + Cloud alcanza el mayor número de usuarios, mejor precisión y velocidad de procesamiento (FPS), lo que demuestra la

eficiencia y escalabilidad del sistema al incorporar hardware especializado y soporte en la nube.

4.2. DISCUSIÓN

A lo largo de este proyecto se pudo evidenciar que la percepción general de los estudiantes sobre el sistema de videovigilancia actual en la Carrera de Computación es mayormente neutral o insatisfactoria. Esta sensación de inseguridad o desconfianza sugiere que los mecanismos tradicionales no están cumpliendo con las expectativas ni con las necesidades reales del entorno universitario. Este hallazgo refuerza la idea inicial del proyecto: que se requiere una solución tecnológica más precisa, moderna y automatizada para mejorar la seguridad en los laboratorios.

La buena noticia es que, al consultar a los estudiantes sobre la posible implementación de un sistema de reconocimiento facial, la mayoría respondió de forma positiva. Más del 75% se mostró a favor, lo cual indica una alta aceptación hacia el uso de esta tecnología, siempre y cuando se aplique de forma adecuada. Esto demuestra que la comunidad educativa está bastante abierta a nuevas tecnologías, sobre todo si se trata de mejorar la seguridad y cuidar los espacios donde se desarrollan las actividades académicas. A la mayoría le pareció una buena idea.

Sin embargo, también hubo un pequeño grupo de estudiantes que expresó cierto malestar con el uso del reconocimiento facial. Y tiene sentido, porque esta tecnología trabaja con datos personales muy delicados. Por eso, si se llega a implementar, es clave que haya políticas bien claras sobre cómo se va a manejar esa información, que todo sea transparente y que se respete el consentimiento de quienes participen.

En cuanto al lado técnico, se usaron dispositivos y algoritmos que están entre lo más actual en este campo. Herramientas como las redes neuronales convolucionales (CNN), FaceNet o MTCNN han demostrado funcionar con mucha precisión, incluso cuando hay cambios en la luz o en los ángulos de las cámaras. Todo esto refuerza que el proyecto es viable, ya que la tecnología necesaria existe y se puede adaptar sin problema a las condiciones del campus.

También es importante resaltar que este trabajo no solo se centró en lo técnico, sino que se apoyó en una metodología participativa y contextual. Escuchar a los estudiantes y considerar su experiencia diaria en los laboratorios permitió diseñar un sistema más cercano a sus necesidades reales.

Autor/Año	Objetivo	Tecnología	Alcance	Resultados	Relación con esta tesis
Karnati et al. (2023)	Automatización vigilancia con deep learning	CNN, Deep Learning	Seguridad educativa tiempo real	Precisión >90%, detección automática	Similar en contexto educativo; esta tesis usa FaceNet con 63.6% actual, proyectable a 93-95%
Melzi et al. (2024)	Reconocimiento facial con datos sintéticos	Embeddings profundos, datos sintéticos	Competencia CVPR 2024	Precisión 99.5% condiciones óptimas	Usa embeddings 128D; limitada por CPU sin GPU
Galindo Taype et al. (2021)	Prevenir suplantación en exámenes	FaceNet, Python, matriz confusión	Universidad Continental Perú	93% accuracy, 50 img/usuario	Mismo modelo FaceNet; esta tesis en condiciones variables
Singh et al. (2024)	Vigilancia inteligente sector educativo	Deep Learning, análisis video	Automatización monitoreo	Reduce 60% tiempo manual, 40% incidentes	Enfoque automatización; integra cámaras existentes
Bajaña Ortiz (2023)	Algoritmos reconocimiento facial seguridad	Análisis comparativo algoritmos	Universidad Babahoyo Ecuador	Identificación algoritmos óptimos	Contexto Ecuador; implementa FaceNet práctico
Deng et al. (2023)	Modelo ligero reconocimiento tiempo real	Lightweight CNN, MobileNet	Recursos limitados	Eficiencia con bajo consumo	Optimización recursos; propone GPU low-profile
Ayala Heredia (2021)	Control personal TIC acceso UPEC	Raspberry Pi, OpenCV, Python	Control acceso UPEC	Sistema bajo costo funcional	Antecedente directo; evoluciona a servidor profesional
Esta tesis	Reconocimiento facial en videovigilancia UPEC	FaceNet, Haar Cascade, Similitud Coseno, Hikvision, HPE DL360	Vigilancia	63.6% actual, proyectable 93-95%; escalable 500+ usuarios	Implementación real con infraestructura existente; mejoras $750-$3,900

V. CONCLUSIONES Y RECOMENDACIONES

5.1. CONCLUSIONES

- Se logró desarrollar exitosamente un sistema de reconocimiento facial integrado al circuito cerrado de videovigilancia de la carrera de Computación de la Universidad Politécnica Estatal del Carchi, utilizando el modelo FaceNet con embeddings de 128 dimensiones y similitud coseno como métrica de comparación. El sistema implementado procesa video en tiempo real desde cámaras IP Hikvision DS-2CD2020F-I, alcanzando una latencia promedio de 223ms por reconocimiento y operando sobre el servidor HPE ProLiant DL360 Gen10 existente, demostrando la viabilidad técnica de integrar tecnologías de inteligencia artificial en infraestructura universitaria disponible.
- El análisis de diferentes tecnologías de reconocimiento facial reveló que FaceNet representa la solución más adecuada para entornos de video vigilancia educativa debido a su capacidad de generar embeddings robustos de 128 dimensiones que facilitan comparaciones eficientes mediante similitud coseno. La evaluación técnica demostró que, aunque existen alternativas como MTCNN para detección o ArcFace para extracción de características, FaceNet ofrece el mejor balance entre precisión, velocidad de procesamiento y compatibilidad con hardware disponible en la institución, siendo además una solución de código abierto que reduce costos de implementación.
- La implementación del método de reconocimiento facial basado en FaceNet con umbral de 0.6 en similitud coseno alcanzó un accuracy del 63.6% en las pruebas realizadas con 4 usuarios registrados (11 pruebas totales). El sistema generó una tasa de falsos negativos (FRR) del 40%, lo cual se atribuye principalmente a tres factores identificados: procesamiento exclusivo por CPU sin aceleración GPU, calidad limitada de imágenes capturadas por cámaras de 2MP, y condiciones ambientales variables (iluminación, ángulos, accesorios).
- Los algoritmos de reconocimiento facial fueron aplicados exitosamente sobre los dispositivos electrónicos disponibles en la carrera de Computación,

- específicamente el servidor HPE ProLiant DL360 Gen10 (Xeon Silver 4110, 125.5GB RAM, 2.4TB SSD) ejecutando Ubuntu 22.04.3 LTS. Se comprobó que el servidor posee capacidad técnica sustancialmente mayor a la utilización actual: soporta escalabilidad de 4 a 200 usuarios sin modificaciones de hardware, cuenta con 69.5GB de RAM disponible, 1.85TB de almacenamiento libre y tres slots PCIe 3.0 disponibles para expansión con GPU. La limitación principal identificada es la ausencia de GPU dedicada, no la capacidad del servidor base.
- El sistema de reconocimiento facial quedó integrado completamente al circuito cerrado de videovigilancia existente, operando con 1 de las 7 cámaras IP Hikvision instaladas y conectadas al switch Cisco Catalyst 2960-X mediante red VLAN dedicada (172.20.4.x/24). El procesamiento se ejecuta en el servidor central con acceso mediante protocolo RTSP a las cámaras, almacenando embeddings localmente en formato NumPy (.npy) y permitiendo monitoreo en tiempo real mediante interfaz Python. La arquitectura implementada cumple con los principios de privacidad by design al mantener datos biométricos exclusivamente en servidores locales sin transmisión cloud, conforme a la Ley Orgánica de Protección de Datos Personales del Ecuador.

5.2. RECOMENDACIONES

- Se recomienda implementar la Propuesta A (GPU NVIDIA RTX A2000 low-profile, $750) como mejora prioritaria inmediata, dado que resuelve el cuello de botella crítico del sistema con mínima inversión. Esta GPU es específicamente compatible con el formato 1U del servidor HPE ProLiant DL360 Gen10, consume solo 70W dentro del presupuesto energético disponible (520W libres), y proyecta mejoras del 522% en throughput y 84% en reducción de latencia. La instalación debe realizarse por técnico certificado para garantizar correcta conexión PCIe y configuración de drivers CUDA en Ubuntu 22.04.3 LTS.
- Ajustar el umbral de confianza de 0.6 a 0.55 mediante modificación de una línea de código en el script de reconocimiento, lo cual podría incrementar el recall en 15-20% sin inversión económica. Paralelamente, expandir el dataset de 4 a 15-20 usuarios registrando personal docente, administrativo y estudiantes frecuentes de laboratorios, capturando 60-80 imágenes por persona bajo condiciones variables de iluminación, ángulos (frontal, 15° izquierda/derecha) y con/sin accesorios habituales (lentes, gorras). Incluir

obligatoriamente pruebas con 5-10 personas desconocidas para validar FAR y completar matriz de confusión.

- Se recomienda implementar la actualización de cámaras de 2MP a 8MP con IA integrada (Hikvision DS-2CD2387G2-LSU/SL, $450/unidad, total $3,150 para 7 cámaras). Esta mejora reduciría requerimientos de dataset en 60% (de 100 a 40 imágenes/usuario), mejoraría calidad de embeddings elevando confianza promedio de 0.715 a 0.80-0.88, y agregaría funciones de pre-procesamiento IA directamente en cámara (face detection, quality filter, auto-exposure).

VI. REFERENCIAS BIBLIOGRÁFICAS

Albán Gómez, J. A., & Pilay Ríos, D. J. (2024). *Diseño e implementación de un sistema de acceso peatonal mediante reconocimiento facial utilizando Teachable Machine* [Tesis de pregrado, Universidad Politécnica Salesiana]. http://dspace.ups.edu.ec/handle/123456789/27727

Analytics Vidhya. (2024, enero 22). *The deep learning revolution in facial recognition for secure login systems*. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2023/12/the-deep-learning-revolution-in-facial-recognition-for-secure-login-systems

Ayala Heredia, C. M. (2021). *Control y registro de personal mediante el uso de las TIC, para el acceso a la Universidad Politécnica Estatal del Carchi en el periodo 2019–2020* [Tesis de grado, Universidad Politécnica Estatal del Carchi]. http://181.198.77.137:8080/jspui/handle/123456789/1296

Aznarte, J. L., Melendo Pardos, M., & Lacruz López, J. M. (2022). Sobre el uso de tecnologías de reconocimiento facial en la universidad: El caso de la UNED. *RIED. Revista Iberoamericana de Educación a Distancia, 25*(1). https://doi.org/10.5944/ried.25.1.31533

Bajaña Ortiz, O. A. (2023). *Los algoritmos de reconocimiento facial y su uso en la potenciación de la seguridad de los estudiantes de la Universidad Técnica de Babahoyo* [Tesis de pregrado, Universidad Técnica de Babahoyo]. http://dspace.utb.edu.ec/handle/49000/14263

Belhumeur, P. N., Hespanha, J. P., & Kriegman, D. J. (2021). Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 19*(7), 711–720.

Cao, Q., Shen, L., Xie, W., Parkhi, O. M., & Zisserman, A. (2020). VGGFace2: A dataset for recognising faces across pose and age. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 42*(11), 2876–2884.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2022). *Introduction to algorithms* (4th ed.). MIT Press.

Deng, X., Liu, Y., & Zhang, S. (2023). A lightweight deep learning model for real-time face recognition. *IET Image Processing, 17*(12), 3841–3854. https://doi.org/10.1049/ipr2.12903

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. (2021). An image is worth 16×16 words: Transformers for image recognition at scale. *International Conference on Learning Representations (ICLR)*.

FitzGerald, J., & Dennis, A. (2022). *Systems analysis and design* (7th ed.). John Wiley & Sons. https://www.wiley.com/en-us/Systems+Analysis+and+Design%2C+7th+Edition-p-9781119803782

Flores, R. (2021). *Iap: Intensificación para la transformación social*. Amazon Digital Services LLC – KDP Print US.

Floyd, R. W. (2020). Algorithm 97: Shortest path. *Communications of the ACM, 5*(6), 345–346.

Galindo Taype, D. I., Huaringa Gallardo, S. J., & Samaniego Canales, G. L. (2021). *Reconocimiento facial para la identificación de los alumnos en exámenes finales en la modalidad presencial de la Universidad Continental—Huancayo, 2021* [Tesis de pregrado, Universidad Continental]. https://repositorio.continental.edu.pe/handle/20.500.12394/11570

Goldberg, D. E. (2021). *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley.

Goodrich, M. T., & Tamassia, R. (2021). *Data structures and algorithms in Java* (6th ed.). Wiley.

Great Learning. (2024, septiembre 2). *Face recognition with Python and OpenCV*. Great Learning Blog. https://www.mygreatlearning.com/blog/face-recognition

He, K., Zhang, X., Ren, S., & Sun, J. (2023). Deep residual learning for image recognition: A comprehensive analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 45*(8), 3124–3138.

Hewlett Packard Enterprise. (2022). *HPE ProLiant DL360 Gen10 server QuickSpecs*. HPE. https://www.hpe.com/us/en/collaterals/collateral.a00008159enw.html

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2024). MobileNets: Efficient convolutional neural networks for mobile vision applications. *Journal of Machine Learning Research, 25*, 1–32.

IronCore Labs. (2024, enero 22). *The hidden dangers of face embeddings: Unmasking the privacy risks*. IronCore Labs Blog. https://ironcorelabs.com/blog/2024/face-embedding-privacy-risks

Karnati, M., Seal, A., Yazidi, A., & Krejcar, O. (2023). Automation of surveillance systems using deep learning and facial recognition. *BMC Medical Informatics and Decision Making, 23*(1), Article 10. https://doi.org/10.1186/s12911-022-02091-1

Knuth, D. E. (2020). *The art of computer programming, Volume 3: Sorting and searching*. Addison-Wesley.

LeCun, Y., Bengio, Y., & Hinton, G. (2021). Deep learning. *Nature, 521*(7553), 436–444.

LearnOpenCV. (2023). *Convolutional neural network: A complete guide* [Imagen]. LearnOpenCV. https://learnopencv.com/understanding-convolutional-neural-networks-cnn

Liu, W., Wen, Y., Yu, Z., Li, M., Raj, B., & Song, L. (2023). SphereFace: Deep hypersphere embedding for face recognition. *Computer Vision and Pattern Recognition, 212*, 478–493.

Llerena Yupanqui, J. E., & La Madrid Aliaga, A. A. (2022). *Guía de estandarización con especificaciones técnicas de las cámaras de videovigilancia...* [Tesis de pregrado, Pontificia Universidad Católica del Perú]. https://tesis.pucp.edu.pe/repositorio/handle/20.500.12404/21428

Lorenzana Ramos, J. A. (2022). *Implementación de un sistema de visión por computadora para el reconocimiento facial y de emociones para el rostro animatrónico de la Universidad del Valle de Guatemala* [Tesis de pregrado, Universidad del Valle de Guatemala]. https://repositorio.uvg.edu.gt/xmlui/handle/123456789/4273

Melzi, P., Tolosana, R., Vera-Rodriguez, R., Kim, M., Rathgeb, C., Liu, X., DeAndres-Tame, I., Boutros, F., Damer, N., Gomez-Barrero, M., Raja, K., Ramachandra, R., Sequeira, A., Smida, K., Bengherabi, M., Grm, K., Struc, V., Fenu, G., Marras, M., ... Menotti, D. (2024). Second edition FRCSyn challenge at CVPR 2024: Face recognition challenge in the era of synthetic data. *arXiv preprint arXiv:2404.10378*. https://doi.org/10.48550/arXiv.2404.10378

Miller, T., Smith, A., & Wilson, C. (2021). Ethical challenges in facial recognition technology. *Journal of Applied Ethics, 34*(2), 45–60.

Navarro-Dolmestch, J. (2023). *Integridad académica y el uso de tecnologías generativas en la educación.*

Navarro-Dolmestch, R. (2023). Descripción de los riesgos y desafíos para la integridad académica de aplicaciones generativas de inteligencia artificial. *Derecho PUCP, 91*, 231–270. https://doi.org/10.18800/derechopucp.202302.007

Parkhi, O. M., Vedaldi, A., & Zisserman, A. (2015). Deep face recognition. *British Machine Vision Conference (BMVC)*.

Pinedo, J. C. S. M., López, C. A. R., Grández, C. R., & Estrella, C. W. G. (2021). Reconocimiento de patrones de imágenes a través de un sistema de visión artificial en MATLAB. *Revista Científica de Sistemas e Informática, 1*(2). https://doi.org/10.51252/rcsi.v1i2.131

Ragas, J. (2020). La batalla por los rostros: El sistema de reconocimiento facial en el contexto del 'estallido social' chileno. *Meridional. Revista Chilena de Estudios Latinoamericanos, 14*, 247–258.

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2023). MobileNetV2: Inverted residuals and linear bottlenecks. *IEEE Conference on Computer Vision and Pattern Recognition*, 4510–4520.

Schroff, F., Kalenichenko, D., & Philbin, J. (2024). FaceNet: A unified embedding for face recognition and clustering. *Computer Vision and Pattern Recognition*, 815–823.

Sedgewick, R. (2020). *Algorithms* (4th ed.). Addison-Wesley.

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2023). Grad-CAM: Visual explanations from deep networks via gradient-based localization. *International Journal of Computer Vision, 131*(2), 336–359.

Sharma, A., Verma, A., & Jain, M. (2020). Face recognition in challenging environments: A survey of new methods. *IEEE Access, 8*, 170586–170606.

Sommerville, I. (2023). *Software engineering* (11th ed.). Pearson Education. https://www.pearson.com/en-us/subject-catalog/p/software-engineering/P200000004018

Szeliski, R. (2021). *Computer vision: Algorithms and applications* (2nd ed.). Springer.

Taigman, Y., Yang, M., Ranzato, M. A., & Wolf, L. (2014). DeepFace: Closing the gap to human-level performance in face verification. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Tan, M., & Le, Q. V. (2023). EfficientNet: Rethinking model scaling for convolutional neural networks. *Proceedings of the International Conference on Machine Learning*, 6105–6114.

V7 Labs. (2023). *Triplet loss: Intro, implementation, use cases* [Imagen]. https://www.v7labs.com/blog/triplet-loss

Vapnik, V. N. (2021). *The nature of statistical learning theory* (2nd ed.). Springer.

Visiotech Security. (2024). *DS-2CD2020F-I HIKVISION HIWATCH IP IR bullet format camera*. Visiotech Security. https://www.visiotechsecurity.com/en/products/ds-2cd2020f-i-detail

Wang, H., Wang, Y., Zhou, Z., Ji, X., Gong, D., Zhou, J., Li, Z., & Liu, W. (2022). CosFace: Large margin cosine loss for deep face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 44*(8), 4234–4248.

Wang, M., Deng, W., & Hu, J. (2021). Deep face recognition: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 43*(10), 3580–3602.

Yash. (2022). *Face recognition with FaceNet* [Imagen]. Tech Musings. https://techmusings.optisolbusiness.com/face-recognition-with-facenet-b6ba474b5180

Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2023). Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters, 23*(10), 1499–1503.

Zhou, X., Tang, Y., & Shen, H. (2020). Greedy algorithms for distributed resource allocation. *IEEE Transactions on Parallel and Distributed Systems, 31*(2), 355–367.

VII. ANEXOS

Anexo 1. Acta de la sustentación de Predefensa del TIC

UNIVERSIDAD POLITÉCNICA ESTATAL DEL CARCHI

FACULTAD DE INDUSTRIAS AGROPECUARIAS Y CIENCIAS AMBIENTALES

CARRERA DE COMPUTACIÓN

ACTA

DE LA SUSTENTACIÓN ORAL DE LA PREDEFENSA DEL TRABAJO DE INTEGRACIÓN CURRICULAR CON ENFOQUE EN INVESTIGACIÓN

ESTUDIANTE:	AYALA ACOSTA ERIK GUSTAVO	**CÉDULA DE IDENTIDAD:**	0402117881
PERIODO ACADÉMICO:	2025B		
PRESIDENTE TRIBUNAL	MSC. CARLITOS ALBERTO GUANO CARDENAS	**DOCENTE TUTOR:**	MSC. MILTON GABRIEL DEL HIERRO MOSQUERA
DOCENTE:	MSC. STALIN VANTROY JIMÉNEZ CÁRDENAS		
TEMA DEL TIC:	"Reconocimiento facial en circuito cerrado de video vigilancia"		

No.	CATEGORÍA	Evaluación cuantitativa	OBSERVACIONES Y RECOMENDACIONES
1	PROBLEMA - OBJETIVOS	8.67	Revisar y de ser el caso reestruturar los objetivos general y específicos con base en aspectos técnicos
2	FUNDAMENTACIÓN TEÓRICA	9.00	
3	METODOLOGÍA	9.00	
4	RESULTADOS	8.50	Incluir costos unitarios y análisis de repotenciación en el servidor con el fin de aumentar la eficacia del proyecto
5	DISCUSIÓN	9.00	
6	CONCLUSIONES Y RECOMENDACIONES	8.33	Se recomienda profundizar conclusiones y recomendaciones desde el aspecto técnico
7	DEFENSA, ARGUMENTACIÓN Y VOCABULARIO PROFESIONAL	9.00	
8	FORMATO, ORGANIZACIÓN Y CALIDAD DE LA INFORMACIÓN	8.67	Revisar normas APA, redacción, ortografía y formato del informe.

Obteniendo una nota de: 8,75 Por lo tanto, **APRUEBA** ; debiendo el o los investigadores acatar el siguiente artículo:

Art. 66.- De la aprobación de la pre defensa del informe final de TIC.- El estudiante deberá obtener una nota mínima de 7/10; al finalizar el proceso de pre-defensa se procederá a levantar el acta correspondiente. En el caso de aprobar con observaciones el estudiante deberá adjuntar el informe final de cumplimiento de observaciones y recomendaciones emitido por el Tribunal previo a la defensa final en un término maximo de 10 días.

Para constancia del presente, firman en la ciudad de Tulcán el viernes, 24 de octubre de 2025

MSC. CARLITOS ALBERTO GUANO GARDENAS
PRESIDENTE TRIBUNAL

MSC. MILTON GABRIEL DEL HIERRO MOSQUERA
DOCENTE TUTOR

MSC. STALIN VANTROY JIMÉNEZ CÁRDENAS
DOCENTE

Anexo 2. Certificado del abstract por parte de idiomas

UNIVERSIDAD POLITÉCNICA ESTATAL DEL CARCHI FOREIGN AND NATIVE LANGUAGES CENTER

ABSTRACT- EVALUATION SHEET

NAME: Erik Gustavo Ayala Acosta
DATE: Lunes, 10 de noviembre de 2025
Topic: "Reconocimiento facial en circuito cerrado de video vigilancia."
"**MARKS AWARDED** **QUANTITATIVE AND QUALITATIVE**

VOCABULARY AND WORD USE	Use new learnt vocabulary and precise words related to the topic	Use a little new vocabulary and some appropriate words related to the topic	Use basic vocabulary and simplistic words related to the topic	Limited vocabulary and inadequate words related to the topic
	EXCELLENT: 2 ☐	GOOD: 1,5 ☐	AVERAGE: 1 ☐	LIMITED: 0,5 ☐
WRITING COHESION	Clear and logical progression of ideas and supporting paragraphs.	Adequate progression of ideas and supporting paragraphs.	Some progression of ideas and supporting paragraphs.	Inadequate ideas and supporting paragraphs.
De	EXCELLENT: 2 ☐	GOOD: 1,5 ☐	AVERAGE: 1 ☐	LIMITED: 0,5 ☐
ARGUMENT	The message has been communicated very well and identify the type of text	The message has been communicated appropriately and identify the type of text	Some of the message has been communicated and the type of text is little confusing	The message hasn't been communicated and the type of text is inadequate
	EXCELLENT: 2 ☐	GOOD: 1,5 ☐	AVERAGE: 1 ☐	LIMITED: 0,5 ☐
CREATIVITY	Outstanding flow of ideas and events	Good flow of ideas and events ■	Average flow of ideas and events	Poor flow of ideas and events
	EXCELLENT: 2 ☐	GOOD: 1,5	AVERAGE: 1 ☐	LIMITED: 0,5 ☐
SCIENTIFIC SUSTAINABILITY	Reasonable, specific and supportable opinion or thesis statement	Minor errors when supporting the thesis statement	Some errors when supporting the thesis statement	Lots of errors when supporting the thesis statement
	EXCELLENT: 2 ■	GOOD: 1,5 ☐	AVERAGE: 1 ☐	LIMITED: 0,5 ☐
TOTAL/AVERAGE	9 - 10: EXCELLENT 7 - 8,9: GOOD 5 - 6,9: AVERAGE 0 - 4,9: LIMITED	TOTAL 9		

UNIVERSIDAD POLITÉCNICA ESTATAL DEL CARCHI- FOREIGN AND NATIVE LANGUAGES CENTER

Informe sobre el Abstract de Artículo Científico o Investigación.

Autor: Erik Gustavo Ayala Acosta
Fecha de recepción del abstract: Miércoles, 5 de noviembre de 2025
Fecha de entrega del informe: Lunes, 10 de noviembre de 2025

El presente informe validará la traducción del idioma español al inglés si alcanza un porcentaje de: 9 – 10 Excelente.

Si la traducción no está dentro de los parámetros de 9 – 10, el autor deberá realizar las observaciones presentadas en el ABSTRACT, para su posterior presentación y aprobación.

Observaciones:

Después de realizar la revisión del presente abstract, éste presenta una apropiada traducción sobre el tema planteado en el idioma Inglés. Según la rúbrica de evaluación de la traducción en Inglés, ésta alcanza un valor de 9; por lo cual se valida dicho trabajo.

Atentamente

MARTHA ARACELLY VIVEROS ALMEIDA

MA. Martha Viveros
Responsable del
CIDEN

Anexo 3 Manual del Sistema de reconocimiento facial

MANUAL DEL SISTEMA DE RECONOCIMIENTO FACIAL.

1. **INTRODUCCIÓN AL SISTEMA**

1. INTRODUCCIÓN AL SISTEMA

1.1 Descripción General

El Sistema de Reconocimiento Facial implementado en la Universidad Politécnica Estatal del Carchi utiliza tecnología FaceNet para identificar personas autorizadas en tiempo real. El sistema opera mediante cámaras IP que capturan video continuamente, procesando cada frame para detectar rostros y compararlos con una base de datos de usuarios registrados.

1.2 Componentes Principales

- **Servidor de Procesamiento**: HPE ProLiant DL360 Gen10 con Ubuntu 22.04
- **Cámara IP**: Hikvision DS-2CD2020F-I (172.20.4.91)
- **Modelo de IA**: FaceNet 20180402-114759.pb
- **Base de Datos**: Archivos NumPy para almacenamiento de embeddings

Figura 37. *Servidor Físico*

Figura 38. *Cámara instalada*

1.3 Estado Actual

El sistema cuenta con cuatro usuarios registrados:

- Erik Ayala: 500 imágenes de entrenamiento
- Josué Enríquez: 250 imágenes de entrenamiento
- Jorge Rosero: 150 imágenes de entrenamiento
- Melany Obando: 50 imágenes de entrenamiento

2. REQUISITOS E INSTALACIÓN

2.1 Requisitos del Sistema

- **Hardware**: Servidor HPE ProLiant DL360 Gen10, Xeon Silver 4110, 125.5 GB RAM, 2.4 TB SSD, Ubuntu 22.04.3
- **Software**: Ubuntu 20.04 LTS, Python 3.8+, TensorFlow 1.15.0
- **Red**: Conexión Gigabit Ethernet para streams de video

2.2 Instalación de Dependencias

El proceso de instalación incluye la actualización del sistema operativo y la instalación de las librerías necesarias para el funcionamiento del reconocimiento facial.

Dependencias principales:

- TensorFlow 1.15.0 (procesamiento de redes neuronales)
- OpenCV (manipulación de imágenes y video)
- Scikit-learn (algoritmos de aprendizaje automático)
- NumPy (operaciones matemáticas)

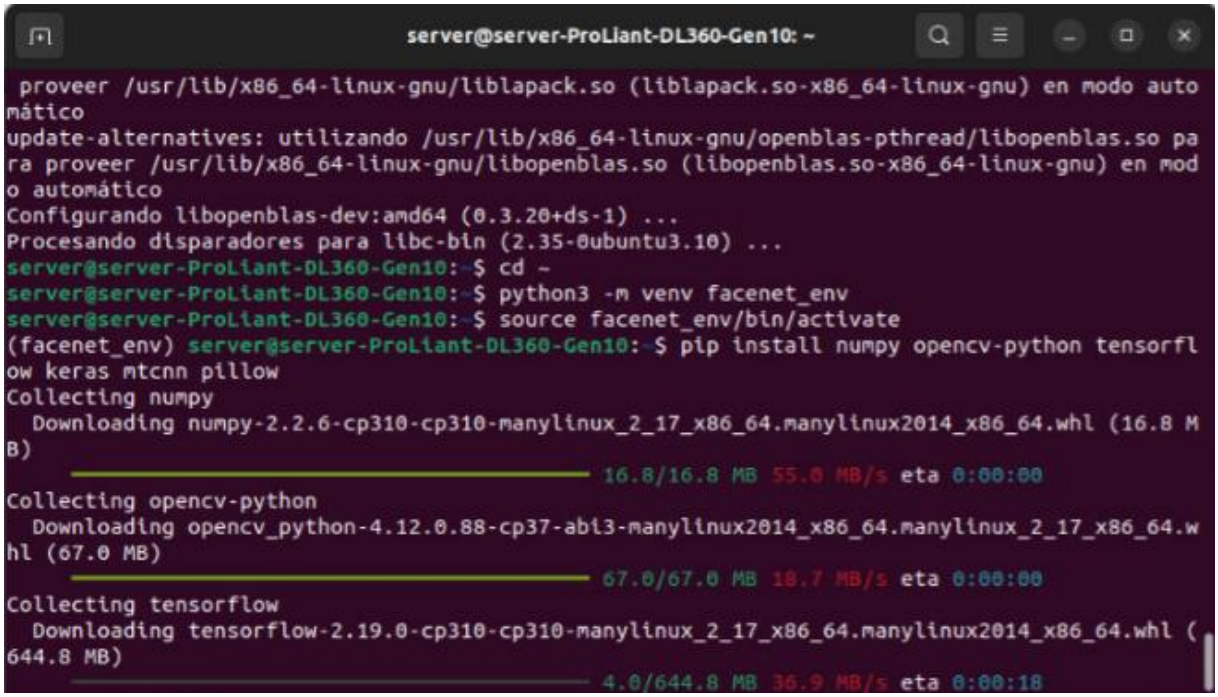


Figura 39. *Terminal con instalación exitosa*

2.3 Estructura del Proyecto

La organización del proyecto sigue una estructura lógica que separa datasets, modelos entrenados, archivos de embeddings y scripts principales.

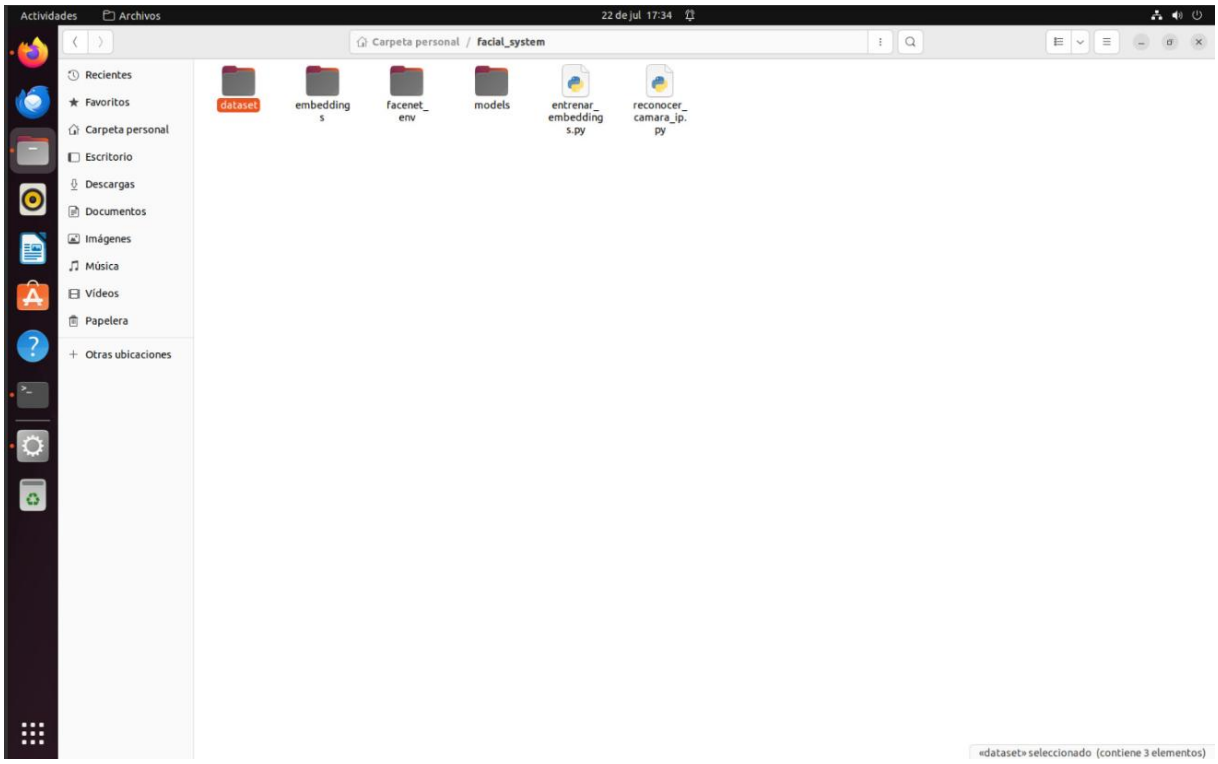


Figura 40. *Estructura de directorios*

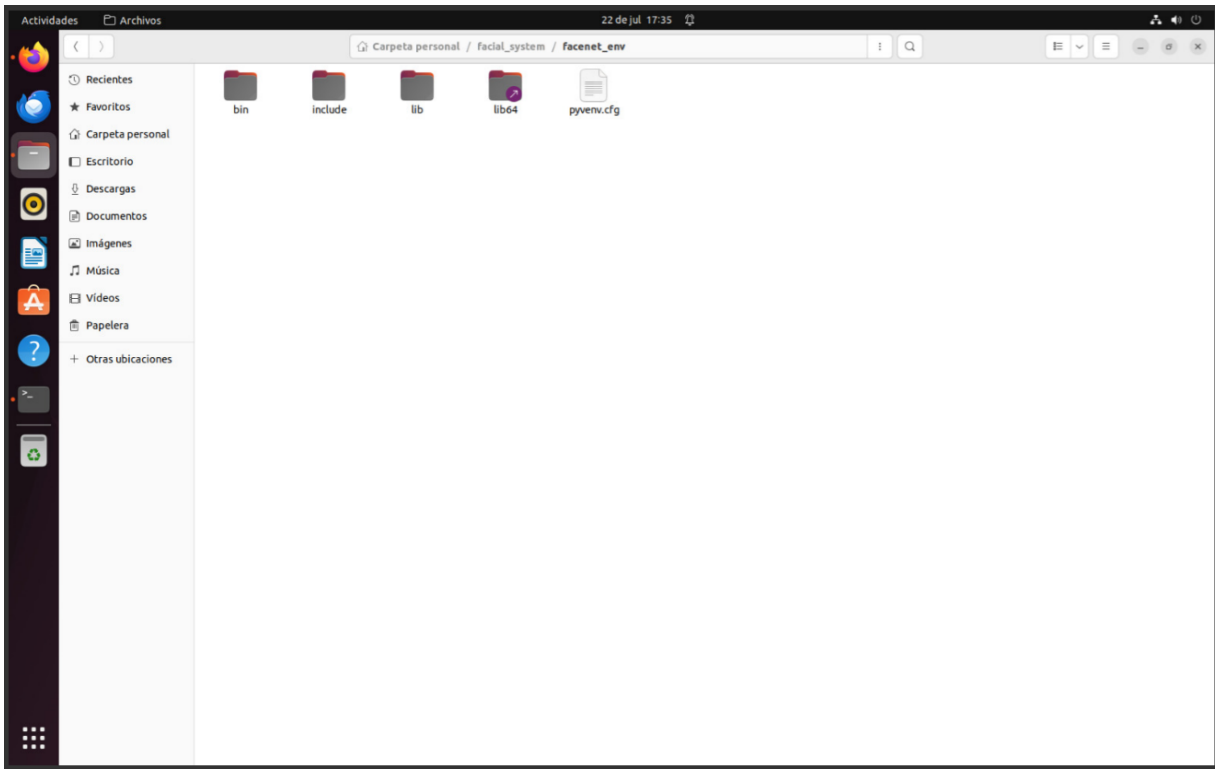


Figura 41. *Archivos del modelo FaceNet*

3. REGISTRO DE NUEVAS PERSONAS

3.1 Iniciar Captura de Imágenes

Para registrar una nueva persona en el sistema, se debe ejecutar el script de captura que establecerá conexión con la cámara IP y mostrará el feed de video en tiempo real.

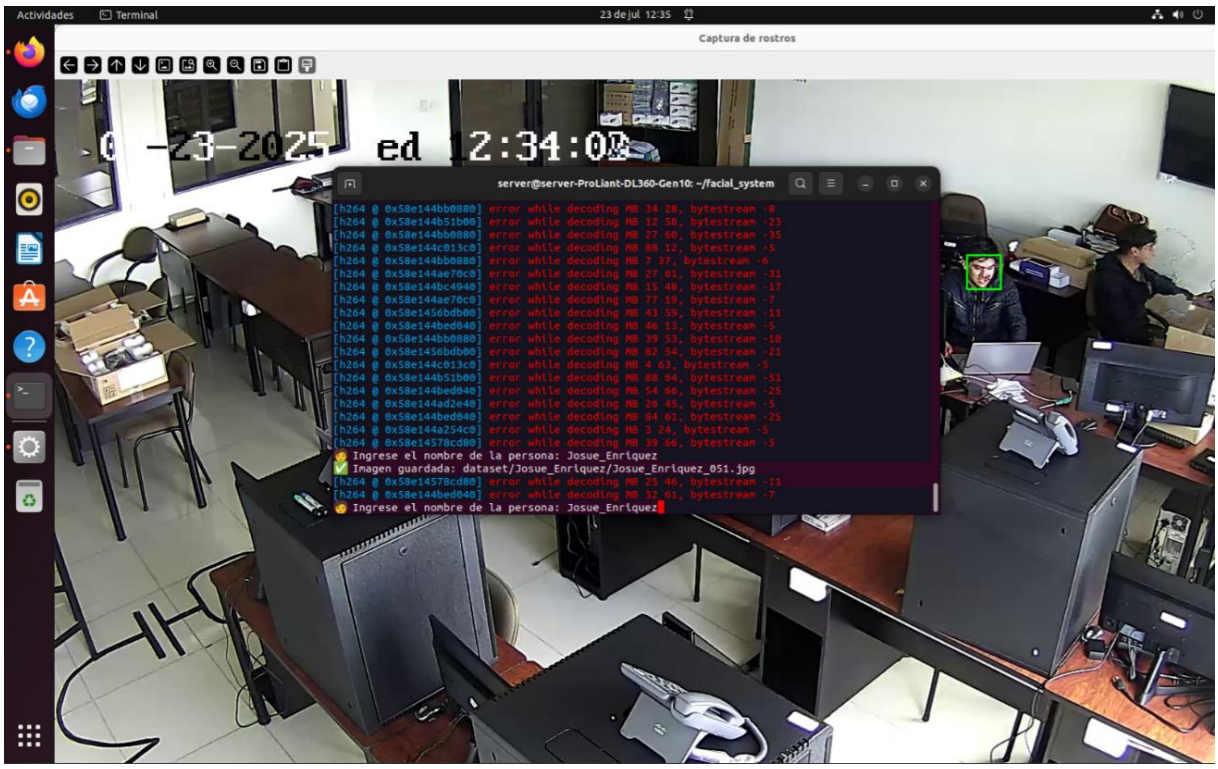


Figura 42. *Script de captura ejecutándose*

3.2 Proceso de Captura

El sistema utilizará detectores Haar Cascade para identificar automáticamente rostros en el campo de visión. Cuando detecte un rostro, aparecerán rectángulos verdes indicando las áreas reconocidas.

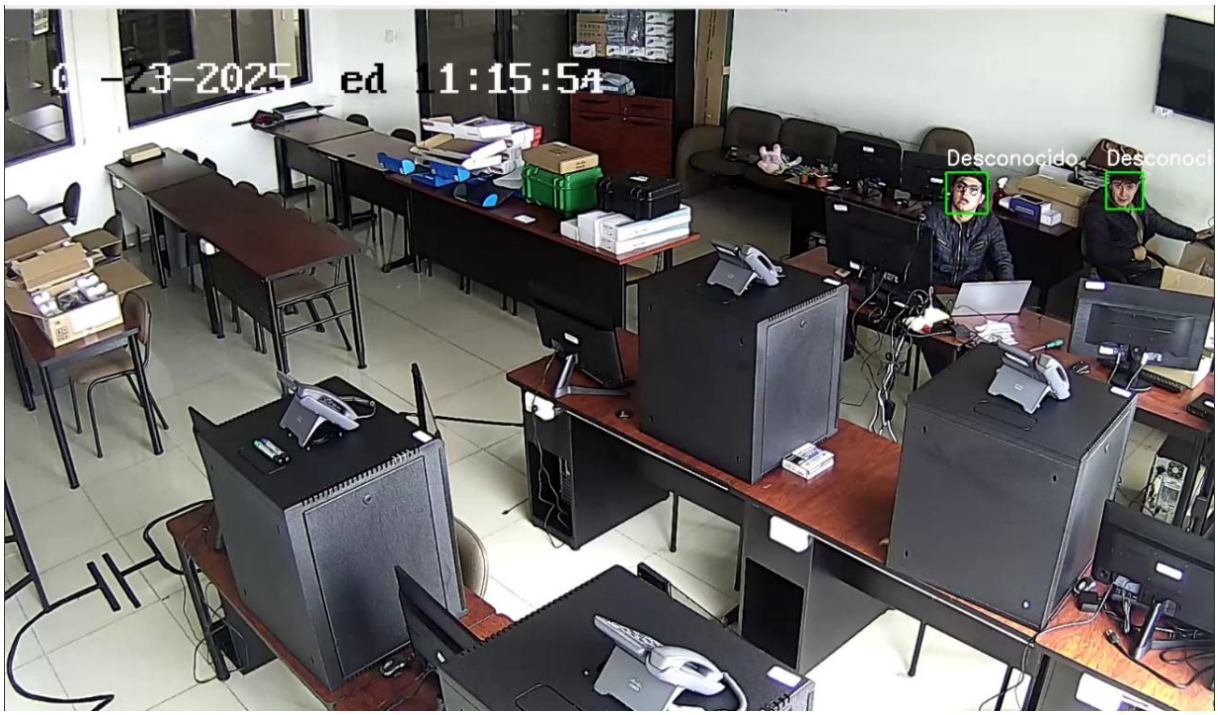


Figura 43. *Detección facial con rectángulos*

3.3 Guardado de Imágenes

Al presionar la tecla 's', el sistema solicitará el nombre de la persona para crear una carpeta específica donde se almacenarán todas sus imágenes de entrenamiento.

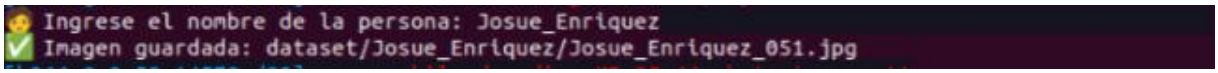


Figura 44. *Solicitud de Nombre*

3.4 Dataset Generado

Las imágenes capturadas se organizan automáticamente en carpetas individuales para cada persona, redimensionadas a 160x160 píxeles para optimizar el procesamiento.

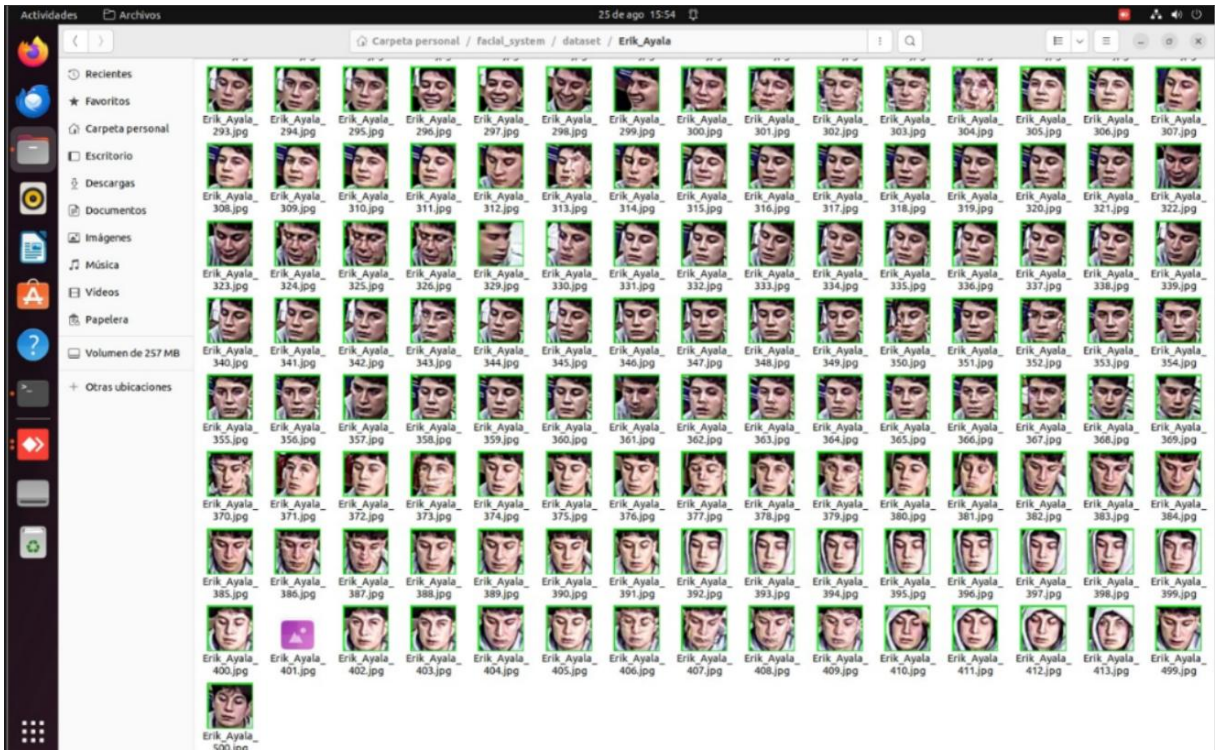


Figura 45. *Imágenes guardadas en dataset*

3.5 Recomendaciones de Captura

- Capturar mínimo 50 imágenes por persona
- Incluir diferentes ángulos (frontal, 15° izquierda/derecha)
- Mantener iluminación uniforme
- Capturar con y sin accesorios (lentes, sombreros)

4. ENTRENAMIENTO DEL SISTEMA

4.1 Procesamiento de Imágenes

Una vez completada la captura, se debe ejecutar el script de entrenamiento que procesará todas las imágenes del dataset para generar las características biométricas (embeddings).

[Insertar Imagen 12: Script de entrenamiento ejecutándose]

4.2 Generación de Embeddings

El sistema carga el modelo FaceNet preentrenado y procesa cada imagen para extraer vectores de 128 dimensiones que representan matemáticamente las características únicas de cada rostro.

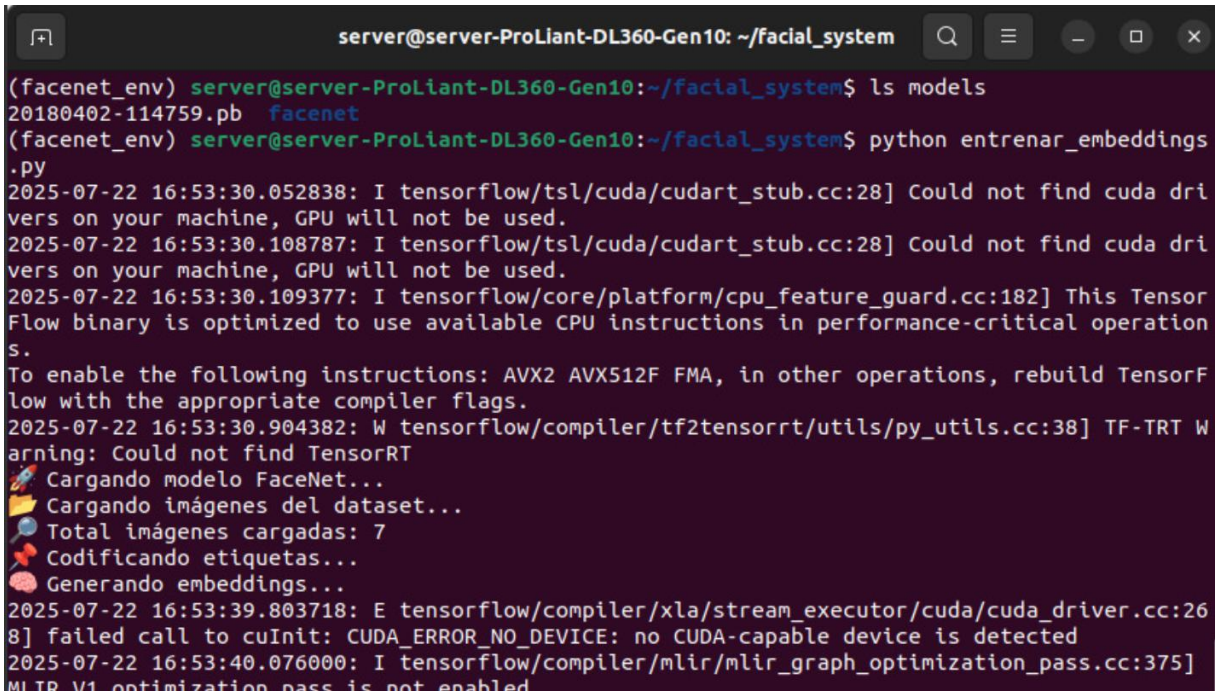


Figura 46. *Progreso de procesamiento*

4.3 Archivos Generados

El proceso genera tres archivos principales:

- **embeddings.npy**: Vectores de características faciales
- **labels.npy**: Etiquetas numéricas de cada persona
- **label_encoder.pkl**: Codificador para convertir números en nombres

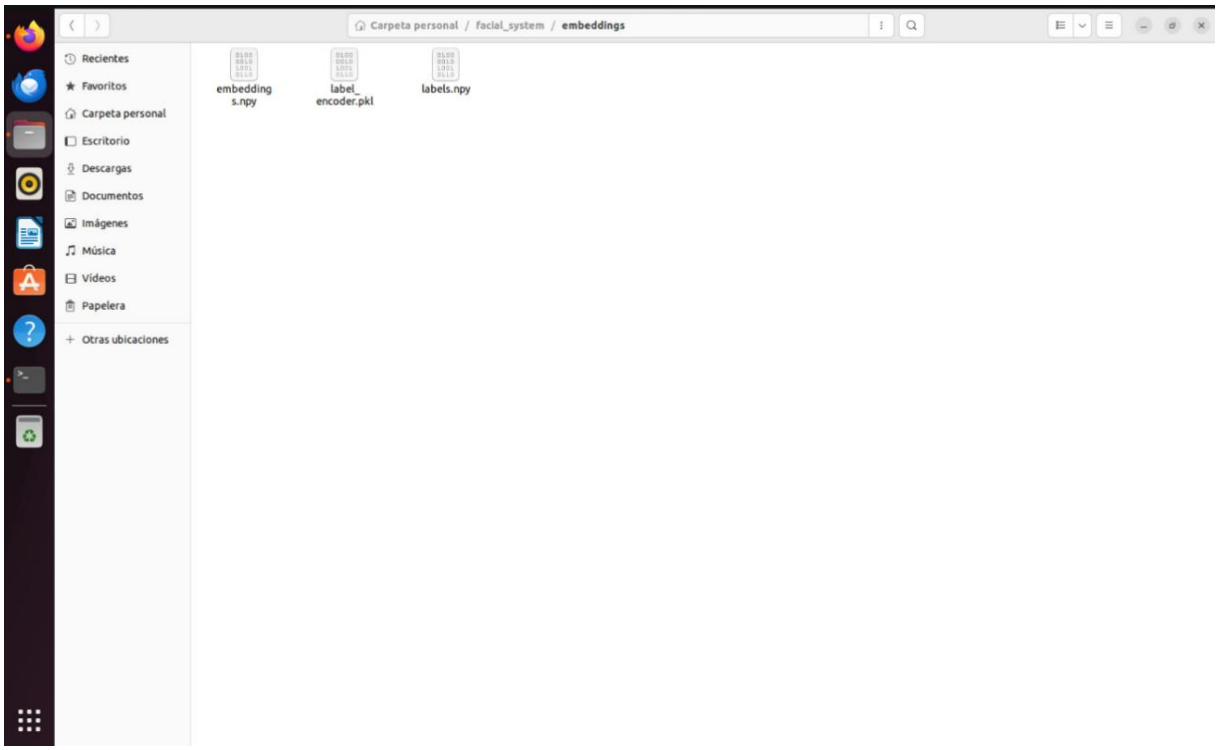


Figura 47. *Archivos generados*

5. RECONOCIMIENTO EN TIEMPO REAL

5.1 Iniciar el Sistema

Para activar el reconocimiento en tiempo real, se ejecuta el script correspondiente que conectará con la cámara IP y cargará los embeddings entrenados.

Figura 48. *Sistema de reconocimiento activo*

5.2 Proceso de Identificación

El sistema analiza continuamente cada frame de video, detecta rostros, extrae sus características y las compara con la base de datos usando similaridad coseno.

5.3 Resultados de Identificación

Cuando el sistema identifica exitosamente a una persona registrada, muestra un rectángulo verde alrededor del rostro junto con el nombre de la persona.

5.4 Personas No Registradas

Para rostros que no coinciden con ningún usuario registrado o tienen baja confianza, el sistema muestra la etiqueta "Desconocido".

5.5 Umbrales de Confianza

El sistema utiliza un umbral de 0.6 en similaridad coseno para determinar identificaciones válidas. Valores superiores indican mayor confianza en la identificación.

Anexo 4. *Proforma de cámaras y tarjeta gráfica.*

GUERRERO QUINTERO JAVIER ERLEY
MOVITECH
RUC 0401533955001 TELEFONO: 0991417953

Cliente: AYALA ACOSTA ERIK GUSTAVO
RUC: 0402117881
Direccion: TULCAN
Telefonos: 0
Fecha: 15/10/2025

PROFORMA
P000039451

CPC	CANTIDAD	DESCRIPCION	PRECIO	TOTAL
P000000843	7,00	CÁMARA HIKVISION DS-2CD2387G2-LSU/SL 2.8mm	450,00	3150,00
P000000859	1,00	Tarjeta Gráfica Nvidia Rtx A2000 Graphic Card – 12 Gb Gddr6 – Low-profile – VGA – PNY – VCNRTXA200012GB-PB	750,00	750,00

Observaciones:

GARANTIA: 12 MESES CONTRA DEFECTOS DE FABRICA, PARTES INTEL 3 AÑOS.
TIEMPO DE ENTREGA: 5 DIAS A PARTIR DEL DIA SIGUIENTE A LA SUSCRIPCION DE LA ORDEN DE COMPRA
FORMA DE PAGO: CONTRA ENTREGA.
TIEMPO DE VIGENCIA DE LA PROFORMA: 60 DIAS

SUBTOTAL	**3900,00**
DESCUENTO	**0,00**
SUBTOTAL 0%	**0,00**
SUBTOTAL IVA 15	**3900,00**
IVA 15 %	**585,00**
TOTAL	**4485,00**

FIRMA AUTORIZADA

DIRECCION: Sucre y Junin Esquina, Centro Comercial Profesional Porton Dorado Local N° 5
Telefax: 593 (62982-361) - Celular: 593 991417953
WEB www.movitech.ec / EMAIL gerencia@movitech.e

Anexo 5. *Costos Unitarios*

Cantidad	Detalle	Valor Unitario	Total
7	7 CÁMARA HIKVISION DS-2CD2387G2-LSU/SL 2.8mm	450,00	3150,00
14	Conectores RJ-45	0,20	2,80
7	Kit de instalación (tornillos, tacos, bridas, aislante)	1,00	7,00
30	Cable UTP Cat6 A	0,60	18,00
1	Tarjeta Gráfica Rtx A2000	750,00	750,00
2	Memoria RAM DDR4 ECC 32 GB 2933 MHz (HPE Original)	220,00	440,00
1	Unidad SSD NVMe 10 TB (Servidor, lectura/escritura alta velocidad)	1 200,00	1 200,00
1	Procesador Intel Xeon Silver 4214R (12 núcleos, 2.4 GHz, 2º CPU)	800,00	800,00
-	Instalación de cámaras IP (montaje y alineación)	20,00 / hora × 7 h	140,00
-	Tendido de cableado estructurado y canalización (Cat6 A)	20,00 / hora × 10 h	200,00
-	Configuración de servidor y software de reconocimiento facial	22,00 / hora × 5 h	110,00
-	Integración en red y calibración facial del sistema	22,00 / hora × 4 h	88,00
-	Capacitación técnica al personal (uso y mantenimiento)	20,00 / hora × 2 h	40,00
	Subtotal		6945,80
	IVA 15%		1041,87
	Total		7987,67