

Sistema de Visión Artificial para la Detección de Rostros Basado en Viola-Jones y Local Binary Patterns en Escenarios Simulados de Derrumbe

Computer Vision System for Face Detection Based on Viola-Jones and Local Binary Patterns in Simulated Collapse Scenarios

Gallegos-Rodríguez, H. A¹ 

Datos de Adscripción:

¹ Gallegos Rodríguez, Hasem Azareel
Tecnológico Nacional de México Campus La Laguna. Torreón, Coahuila.
México.
ROR ID: <https://ror.org/00xbw4a70>
m.hagallegosr@lalaguna.tecnm.mx
ORCID: <https://orcid.org/0009-0005-2509-9574>

Resumen - El presente trabajo describe el desarrollo y evaluación de un sistema de visión artificial para la detección de rostros humanos en escenarios simulados de derrumbe, con el objetivo de apoyar posibles tareas de búsqueda y rescate. El sistema propuesto se basa en la integración de dos técnicas ampliamente utilizadas en el procesamiento de imágenes: el detector Viola-Jones y el descriptor Local Binary Patterns. En el esquema implementado, el detector Viola-Jones realiza una detección inicial de posibles regiones faciales, mientras que el descriptor Local Binary Patterns se emplea para confirmar la presencia del rostro y reducir detecciones incorrectas. La elección de estos métodos se realizó considerando su eficiencia, bajo costo computacional y su amplio uso en aplicaciones de detección facial. El sistema fue desarrollado en Python y con la biblioteca OpenCV, utilizando una cámara IP para la adquisición de imágenes en tiempo real. Las pruebas experimentales se realizaron en distintos ambientes que simulan condiciones asociadas a escenarios de derrumbe, incluyendo variaciones en la iluminación, la presencia de elementos en el entorno y diferentes orientaciones del rostro, así como situaciones en las que únicamente una parte del rostro se encuentra visible. Los resultados obtenidos muestran que la combinación de Viola-Jones y Local Binary Patterns permite realizar detecciones faciales de manera confiable en los escenarios evaluados, alcanzando una efectividad aproximada del 85 %. Estos resultados indican que el sistema desarrollado puede emplearse como una herramienta de apoyo en tareas de búsqueda y rescate.

Palabras Clave: Búsqueda y rescate, Detección de rostros, Escenarios de derrumbe, Patrones binarios locales, Viola-Jones, Visión artificial.

Abstract - This work describes the development and evaluation of a computer vision system for human face detection in simulated collapse scenarios, with the aim of

supporting potential search and rescue operations. The proposed system is based on the integration of two widely used techniques in image processing: the Viola-Jones detector and the Local Binary Patterns descriptor. In the implemented scheme, the Viola-Jones detector performs an initial detection of potential facial regions, while the Local Binary Patterns descriptor is used to confirm the presence of a face and reduce false detections. These methods were selected due to their efficiency, low computational cost, and widespread use in face detection applications. The system was developed using the Python programming language and the OpenCV library, employing an IP camera for real-time image acquisition. Experimental tests were conducted in different environments simulating conditions associated with collapse scenarios, including variations in lighting, the presence of environmental elements, and different face orientations, as well as situations in which only a portion of the face is visible. The results show that the combination of Viola-Jones and Local Binary Patterns enables reliable face detection in the evaluated scenarios, achieving an effectiveness of approximately 85%. These results indicate that the developed system can be used as a support tool in search and rescue operations.

Keywords: Collapse scenarios, Computer Vision, Face Detection, Local Binary Patterns, Search and Rescue, Viola-Jones

I. INTRODUCCIÓN

México se encuentra en una de las regiones con mayor actividad sísmica del mundo debido a su ubicación en el Cinturón de Fuego del Pacífico, donde convergen distintas placas tectónicas. De acuerdo con el Centro Nacional de Prevención de Desastres (CENAPRED, 2019), a lo largo del tiempo, eventos como los terremotos de 1985 y 2017 han evidenciado no solo la vulnerabilidad de la infraestructura, sino también la dificultad para localizar a personas atrapadas tras un colapso estructural. Esta problemática se vuelve especialmente crítica en las primeras horas posteriores al desastre, donde la rapidez en la localización puede marcar la diferencia entre la vida y la muerte

En estos escenarios, el desafío no consiste únicamente en localizar a las víctimas, sino en realizar esta tarea bajo ambientes altamente adversos. Las labores de búsqueda suelen desarrollarse con baja visibilidad, presencia de escombros y estructuras inestables, lo que incrementa el riesgo para los rescatistas y limita la efectividad de los métodos tradicionales (Murphy et al., 2008; Erdelj et al., 2017). Por esta razón, en los últimos años se ha impulsado el desarrollo de tecnologías capaces de complementar estas tareas y mejorar los tiempos de respuesta (Liu et al., 2020; Restas, 2015).

Dentro de estas tecnologías, la visión artificial ha adquirido un papel relevante al facilitar el análisis automático de imágenes en tiempo real. Su aplicación en escenarios de rescate ha demostrado que es posible identificar indicios de presencia humana incluso bajo condiciones complejas, lo que la convierte en una herramienta prometedora para este tipo de aplicaciones (Moreno, 2014).

En particular, la identificación facial se ha consolidado como una estrategia efectiva dentro de la visión por computadora, ya que la región facial contiene patrones suficientemente distintivos para identificar la presencia humana incluso cuando otras partes del cuerpo no son visibles. Métodos clásicos como Viola-Jones, basados en características Haar, han sido ampliamente utilizados por su eficiencia en tiempo real (Viola & Jones, 2001), mientras que descriptores como Local Binary Patterns (LBP) permiten representar la textura facial de forma robusta ante variaciones de iluminación (Adeshina et al., 2021; Ahonen et al., 2006). Estas características resultan especialmente relevantes en escenarios de derrumbe, donde la visibilidad suele ser limitada y el rostro puede encontrarse parcialmente oculto.

En este contexto, distintos trabajos han explorado la combinación de detectores basados en Haar con descriptores de textura como LBP, con el objetivo de mejorar la confiabilidad del sistema y reducir falsas detecciones. Este tipo de arquitecturas híbridas permite aprovechar la rapidez de la localización inicial y, al mismo tiempo, refinar los resultados mediante análisis de patrones locales, lo cual resulta adecuado para condiciones complejas como oclusiones parciales y cambios de iluminación (Xu et al., 2016; AlMousawi et al., 2020; Sutikno et al., 2025).

Por otro lado, aunque los métodos basados en aprendizaje profundo han mostrado resultados sobresalientes en tareas de detección, su implementación en sistemas orientados a búsqueda y rescate puede verse limitada por los requerimientos computacionales y la necesidad de grandes volúmenes de datos para su entrenamiento (Jadeja et al., 2024; Zhao, 2024). En contraste, los métodos clásicos continúan ofreciendo ventajas en términos de desempeño computacional y facilidad de implementación en hardware de bajo consumo (Gerónimo et al., 2010), lo que los hace especialmente adecuados para aplicaciones en campo.

Bajo esta perspectiva, resulta relevante analizar y evaluar propuestas basadas en técnicas clásicas de detección facial que integren técnicas clásicas de detección facial en condiciones

controladas que representen escenarios reales. En este trabajo se propone el desarrollo de un sistema de visión artificial para la detección de rostros humanos basado en la integración del detector Viola-Jones y el descriptor Local Binary Patterns. El sistema es evaluado en escenarios simulados de derrumbe, considerando variaciones en la iluminación, presencia de elementos en el entorno y diferentes configuraciones de visibilidad del rostro.

El objetivo principal es evaluar el desempeño de esta combinación en condiciones adversas, buscando que pueda funcionar como una alternativa viable, eficiente y de bajo costo computacional para su posible aplicación en tareas de búsqueda y rescate.

II. PARTE TÉCNICA DEL ARTÍCULO

2.1 Descripción general del sistema

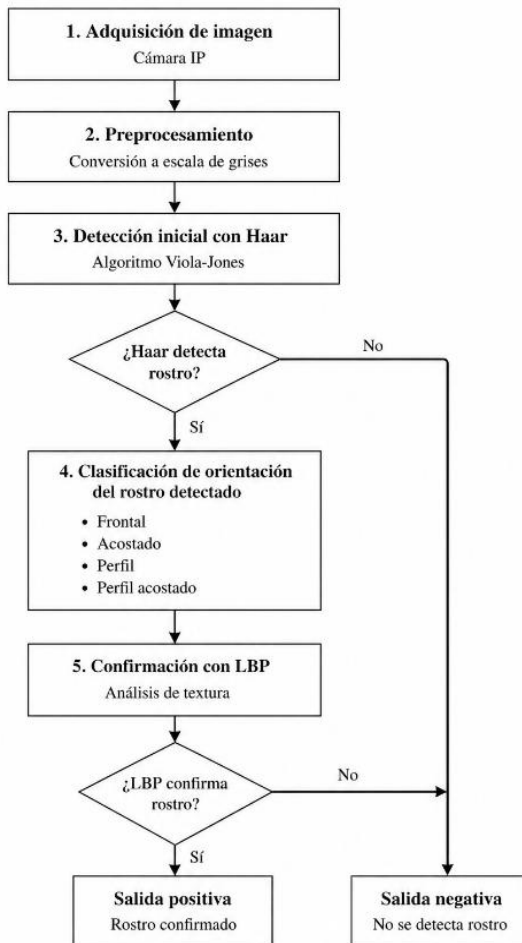
El sistema propuesto tiene como objetivo detectar rostros humanos en tiempo real en condiciones que simulan escenarios de derrumbe, donde factores como la iluminación, los objetos en el entorno o la visibilidad del rostro pueden variar bastante. Para lograrlo, se optó por un enfoque híbrido que combina el detector Viola-Jones con el descriptor Local Binary Patterns (LBP), buscando aprovechar lo mejor de cada método. Por un lado, Viola-Jones permite encontrar rápidamente posibles rostros dentro de la imagen, mientras que LBP ayuda a confirmar si realmente se trata de un rostro, reduciendo así algunas detecciones incorrectas.

De manera general, el sistema sigue una estructura secuencial compuesta por varias etapas: adquisición de imágenes, preprocesamiento, detección inicial y validación de resultados. Este flujo permite que el procesamiento se realice paso a paso, de forma ordenada, desde que se obtiene la imagen hasta que se determina si hay un rostro presente. Cada etapa cumple una función específica y, en conjunto, hace que el sistema sea más estable y confiable.

En la Figura 1 se muestra el diagrama general del sistema propuesto. El flujo comienza con la adquisición de la imagen y su conversión a escala de grises en la etapa de preprocesamiento. Posteriormente, el algoritmo Viola-Jones identifica posibles regiones faciales mediante el uso de cuatro clasificadores entrenados para diferentes orientaciones del rostro: frontal, acostado, perfil y perfil acostado. Si alguno de los clasificadores detecta una región candidata, el sistema procede a una etapa de clasificación de orientación y posteriormente a la confirmación mediante el análisis de textura con el descriptor LBP. Por último, si el patrón es validado, el sistema genera una salida positiva confirmando la detección del rostro.

Figura 1

Diagrama general del sistema de detección propuesto.



2.2 Adquisición y preprocesamiento de imágenes

La adquisición de imágenes se llevó a cabo mediante una cámara IP con capacidad de visión nocturna, capaz de capturar video en tiempo real bajo distintas condiciones de iluminación, incluyendo escenarios con baja o nula luz. Esta característica resultó especialmente útil para simular entornos de derrumbe, donde la iluminación suele ser limitada o irregular. Además, el uso de una cámara de este tipo permitió mantener una calidad de imagen relativamente estable durante todas las pruebas, reduciendo la dependencia de fuentes de luz externas.

Esta fase es fundamental dentro de la arquitectura propuesta, ya que la calidad de las imágenes influye directamente en el desempeño de los algoritmos empleados. Una captura deficiente puede provocar pérdida de información relevante o introducir ruido visual, lo que afecta la localización de características faciales.

Una vez obtenida la captura, se aplicó un preprocesamiento básico basado principalmente en su conversión a escala de grises. Esta transformación permite simplificar la información visual al reducirla a niveles de intensidad, eliminando la información de color que no resulta relevante para los métodos empleados. De esta manera, se disminuye la complejidad computacional y se facilita el procesamiento posterior sin comprometer la detección de rasgos faciales.

2.3 Detección de rostros mediante Viola-Jones

Para la fase inicial de detección de rostros se implementó el algoritmo Viola-Jones, reconocido por su eficiencia en aplicaciones de tiempo real. Este método se basa en el uso de características tipo Haar (Haar-like features) y un clasificador en cascada, permitiendo descartar de forma rápida aquellas regiones de la imagen que no contienen un rostro, reduciendo así el costo computacional del procesamiento (Viola & Jones, 2001). Las características Haar consisten en descriptores que calculan diferencias de intensidad entre regiones rectangulares adyacentes de la imagen, lo que permite representar patrones simples asociados a estructuras faciales.

Dentro de la arquitectura propuesta, el algoritmo tiene como objetivo localizar posibles regiones faciales presentes en la escena. Para ello, analiza múltiples subventanas de la imagen en distintas escalas, lo que facilita la detección de rostros con diferentes tamaños y posiciones (Ambrogio, 2020). No obstante, debido a las condiciones complejas presentes en los escenarios evaluados, pueden generarse falsos positivos ocasionados por objetos o patrones visuales con apariencia similar a un rostro humano.

Por ello, los resultados obtenidos en esta fase son considerados únicamente como regiones candidatas preliminares y no como resultados definitivos. Posteriormente, dichas regiones son sometidas a una etapa adicional de validación, con el propósito de incrementar la precisión y confiabilidad del sistema antes de emitir una decisión final.

Además, para el entrenamiento de los clasificadores empleados en la propuesta, se usó una misma base de datos compuesta por 1125 imágenes faciales, utilizada tanto para el modelo basado en características Haar como para el descriptor LBP. Las imágenes fueron obtenidas considerando diferentes orientaciones del rostro, incluyendo posiciones frontales, acostadas y de perfil, con el propósito de representar condiciones de detección variadas. Asimismo, se incorporaron cambios relacionados con la iluminación, expresiones faciales y presencia parcial de accesorios como lentes o gorras, así como diferencias en el fondo y el entorno de captura.

Cada muestra de la base de datos fue tomada a una resolución de 38×46 píxeles, con el objetivo de reducir el tiempo de procesamiento y facilitar el entrenamiento de los clasificadores faciales sin perder las características principales del rostro.

La base de datos también incluyó rostros capturados desde distintos ángulos y niveles de inclinación, permitiendo disponer de ejemplos faciales con variaciones significativas en la posición de la cabeza. Esto resultó especialmente útil para el entrenamiento de los clasificadores orientados a detectar rostros en posiciones no convencionales, como aquellas que pueden presentarse en escenarios simulados de derrumbe.

En la Figura 2 se muestran algunos ejemplos representativos de las imágenes utilizadas durante el entrenamiento de los clasificadores faciales. En ellas pueden observarse distintas orientaciones del rostro, incluyendo posiciones frontales, laterales y acostadas, así como variaciones en las condiciones visuales y en la presencia de accesorios utilizados durante la captura de las imágenes.

Figura 2

Ejemplos representativos de imágenes de 38x46 utilizadas para el entrenamiento de los clasificadores Haar y LBP.



2.4 Validación de detecciones mediante Local Binary Patterns (LBP)

Con la finalidad de mejorar la precisión del sistema, las regiones identificadas previamente por Viola-Jones son sometidas a un proceso de validación utilizando el descriptor Local Binary Patterns (LBP). Este método analiza la textura local de la imagen mediante la generación de patrones binarios que representan la relación entre un píxel central y los píxeles de su entorno inmediato (Ojala et al., 2002).

La aplicación de LBP permite extraer características relacionadas con la textura facial, incluyendo variaciones de intensidad, contornos y detalles presentes en la superficie del rostro. Gracias a ello, resulta posible diferenciar con mayor eficacia entre un rostro verdadero y resultados incorrectos generados por objetos o estructuras visuales similares. Además, en comparación con otros enfoques más complejos, LBP destaca por su bajo costo computacional y facilidad de implementación, características que

favorecen su uso en aplicaciones de tiempo real (García-Olalla & Alegre, s. f.).

Otro aspecto relevante de este descriptor es su capacidad para mantener un desempeño estable frente a cambios de iluminación. Esta propiedad resulta especialmente importante en escenarios no controlados, donde la calidad visual de las imágenes puede variar considerablemente. Como consecuencia, el sistema puede conservar un funcionamiento más consistente incluso bajo condiciones adversas.

En este contexto, LBP funciona como una fase complementaria dentro del esquema híbrido, contribuyendo a disminuir la presencia de falsos positivos y aumentando la confiabilidad de las detecciones obtenidas.

2.5 Integración del sistema híbrido

La combinación de Viola-Jones y Local Binary Patterns (LBP) constituye la base del sistema híbrido desarrollado en este trabajo. La integración de ambos métodos tiene como finalidad aprovechar las fortalezas particulares de cada técnica dentro de una misma arquitectura de detección facial. En esta estructura, Viola-Jones funciona como la primera etapa del proceso, encargándose de localizar de manera rápida las posibles regiones faciales presentes en la imagen. Posteriormente, las regiones detectadas son evaluadas mediante LBP, el cual analiza las características de textura para verificar si la región corresponde realmente a un rostro humano.

El funcionamiento conjunto de estas técnicas permite establecer un proceso de detección organizado en diferentes etapas de procesamiento. Inicialmente, el detector Viola-Jones realiza un barrido sobre la imagen utilizando características Haar y clasificadores en cascada entrenados previamente para distintas orientaciones faciales. Gracias a ello, el sistema puede identificar regiones candidatas en un tiempo reducido. Sin embargo, debido a que este tipo de detección puede verse afectado por elementos presentes en la escena, se incorpora una etapa adicional de validación basada en LBP.

El descriptor LBP complementa el proceso mediante el análisis de patrones de textura local presentes en las regiones detectadas. Este método genera representaciones binarias a partir de la relación entre los píxeles vecinos, permitiendo describir características faciales de manera más detallada. Como resultado, la etapa de validación ayuda a disminuir detecciones incorrectas ocasionadas por objetos, sombras o variaciones de iluminación dentro del entorno.

La integración secuencial de Viola-Jones y LBP también facilita la distribución de tareas dentro del sistema. Mientras la primera etapa se centra en la localización rápida de posibles rostros, la segunda se orienta a la confirmación de las detecciones obtenidas. Esta organización permite estructurar el flujo de procesamiento de forma más clara, asignando una función específica a cada método dentro de la arquitectura propuesta.

Asimismo, el sistema incorpora clasificadores entrenados para diferentes orientaciones del rostro, incluyendo posiciones frontales, laterales y acostadas. Esto permite ampliar las posibilidades de detección ante escenarios donde la posición facial no se mantiene fija, condición que puede presentarse en ambientes de prueba similares a situaciones reales de derrumbe

En conjunto, el enfoque híbrido permite integrar dos métodos complementarios dentro de una misma arquitectura de visión artificial, combinando rapidez de detección y análisis de textura durante el procesamiento de imágenes. La estructura propuesta establece una base de trabajo orientada a la detección facial en condiciones visuales variables, manteniendo una secuencia de procesamiento adecuada para aplicaciones de monitoreo y análisis en tiempo real.

2.6 Ejemplo de detección del sistema

Con el propósito de mostrar el funcionamiento del sistema propuesto en condiciones similares a un entorno real, en la Figura 3 se muestra una captura obtenida durante la ejecución del sistema en tiempo real mediante una cámara IP. En la imagen se observa a un sujeto con el rostro en posición acostada y parcialmente cubierto por una gorra, mientras el sistema identifica la región facial mediante un recuadro delimitador. Asimismo, puede apreciarse un entorno con presencia de elementos naturales y condiciones de iluminación media durante la captura.

Figura 3

Ejemplo de detección de rostro utilizando el sistema propuesto.



La figura permite visualizar el comportamiento general del sistema durante la etapa de detección facial en tiempo real, así como la forma en que el recuadro generado señala la ubicación del rostro dentro de la escena capturada. Además, el ejemplo mostrado representa una de las condiciones consideradas durante las pruebas experimentales, donde la orientación del rostro y las características del entorno pueden variar.

Cabe señalar que, aunque la figura muestra un caso exitoso, el desempeño global del sistema se analiza de manera cuantitativa en la sección de resultados, donde se consideran múltiples pruebas bajo diferentes condiciones.

2.7 Configuración experimental

Las pruebas del sistema se llevaron a cabo en escenarios simulados de derrumbe, con el objetivo de evaluar su desempeño bajo distintas condiciones que pueden presentarse en situaciones reales de búsqueda y rescate. Para ello, se consideraron principalmente dos factores: las condiciones de iluminación y la presencia de objetos en el entorno.

Con base en estos criterios, se definieron cuatro ambientes de prueba que permiten representar desde condiciones controladas hasta escenarios más complejos. Esta clasificación permite analizar de manera más clara cómo se comporta el sistema ante cambios en el entorno.

En la Tabla 1 se muestran los ambientes utilizados durante la evaluación del sistema.

Tabla 1

Ambientes definidos para la evaluación del sistema.

Escenario	Condición de iluminación	Presencia de objetos	Descripción
1	Baja	No	Escenario con poca iluminación
2	Alta	No	Escenario muy iluminado sin objetos
3	Media	No	Iluminación media sin objetos
4	Media	Si	Iluminación media con presencia de objetos

En cada uno de estos ambientes se evaluaron distintos casos de prueba relacionados con la visibilidad del rostro, considerando variaciones en la orientación, oclusión y nivel de detalle visible. Estos casos permiten analizar el comportamiento del sistema ante diferentes situaciones que pueden afectar la detección facial, como rostros parcialmente cubiertos o en posiciones no convencionales. En la Tabla 2 se muestran los casos de prueba considerados durante la experimentación.

Tabla 2

Casos de prueba evaluados en cada ambiente.

Caso	Descripción
1	Rostro visible y de frente
2	Rostro de perfil parcialmente visible
3	Rostro parcialmente cubierto
4	Solo una parte del rostro visible (ojos o boca)
5	Rostro completamente cubierto
6	Rostro visible desde diferentes ángulos

III. RESULTADOS Y DISCUSIÓN

El desempeño del sistema híbrido propuesto (Viola-Jones + Local Binary Patterns) fue evaluado mediante pruebas en entornos controlados y posteriormente en escenarios no controlados, con el objetivo de analizar su comportamiento ante variaciones de iluminación, oclusión y complejidad del entorno.

3.1 Resultados en entornos controlados

Durante las pruebas, los resultados fueron registrados de manera binaria, considerando detecciones correctas (V) y detecciones incorrectas (F). Se definió como verdadero (V) cuando el sistema detectó correctamente un rostro, y como falso (F) cuando no se detectó un rostro presente en la escena.

Para evaluar el desempeño del sistema, se utilizó la métrica de exactitud, la cual representa la proporción de detecciones correctas respecto al total de pruebas realizadas. En (1) se muestra la expresión utilizada para su cálculo:

$$\text{Exactitud}(\%) = \frac{V}{N} \times 100 \quad (1)$$

donde V representa el número de detecciones correctas y N el total de pruebas realizadas.

En una primera etapa, el sistema fue evaluado en ambientes controlados considerando los distintos escenarios de iluminación y los casos de visibilidad definidos previamente. Para cada combinación de condiciones (ambiente-caso), se realizaron 2 repeticiones experimentales, sumando un total de 48 pruebas.

En la figura 4 se muestran los resultados obtenidos, donde se registran las detecciones correctas (V) y las incorrectas (F) para cada uno de los casos evaluados en los diferentes ambientes.

Figura 4

Resultados de detección del sistema híbrido Haar+LBP.

Ambiente \ Caso	C1	C2	C3	C4	C5	C6
E1 (Oscuro)	V	V	V	F	F	V
E1 (Repetición)	V	V	V	V	F	V
E2 (Muy iluminado)	V	V	V	V	F	V
E2 (Repetición)	V	V	V	F	F	V
E3 (Iluminación media)	V	V	V	F	F	V
E3 (Repetición)	V	V	V	V	F	V
E4 (Media + objetos)	V	V	V	V	F	V
E4 (Repetición)	V	V	V	F	F	V

A partir de los resultados mostrados en la Figura 4, se obtuvo un total de 36 detecciones correctas y 12 incorrectas, lo que corresponde a una exactitud aproximada del 75 %.

Se observa que los casos C1, C2 y C3 presentan un comportamiento consistente en todos los ambientes, manteniendo detecciones correctas independientemente de las condiciones de iluminación. Esto confirma la robustez del sistema ante variaciones lumínicas cuando el rostro es visible.

Por otro lado, los casos C4 y C5 presentan mayores dificultades, especialmente en situaciones de oclusión parcial o total del rostro. En particular, el caso C5 (rostro completamente cubierto) no registró detecciones correctas, lo cual es coherente con la

naturaleza de los algoritmos utilizados, ya que estos dependen de la presencia de características faciales visibles para su funcionamiento.

Asimismo, la repetición de pruebas bajo las mismas condiciones produjo resultados similares, lo que evidencia la consistencia y reproducibilidad del sistema.

3.2 Ajustes del sistema

Con base en los resultados obtenidos en las primeras pruebas, se realizaron diversos ajustes orientados a mejorar el desempeño del sistema en tiempo real. Estos cambios se enfocaron principalmente en reducir la carga computacional y lograr una ejecución más estable, sin comprometer de manera importante la capacidad de detección.

Uno de los ajustes más relevantes fue la reducción de la resolución de procesamiento a 320 × 240 píxeles. Inicialmente, el sistema operaba a resolución Full HD (1080p), correspondiente a la capacidad nativa de la cámara IP utilizada. Sin embargo, se observó que trabajar con esta resolución generaba un procesamiento lento, interrupciones en el flujo de video e incluso cierres inesperados del sistema. Al reducir la resolución, se logró disminuir significativamente el tiempo de procesamiento por cuadro, lo que permitió obtener un flujo de video más fluido y estable durante la ejecución.

De manera complementaria, se limitó el flujo de video a 10 cuadros por segundo (FPS). Aunque este ajuste se realizó inicialmente como una prueba para mejorar el rendimiento general, se observó que contribuía a reducir la carga sobre el sistema y a mantener una ejecución más constante. En la práctica, esta configuración resultó suficiente para conservar una detección funcional en tiempo real, evitando sobrecargas innecesarias sin afectar de forma notable la respuesta del sistema.

Asimismo, se realizó un ajuste en el umbral de brillo con el objetivo de mantener una respuesta más consistente ante variaciones de iluminación. Este parámetro permitió mejorar la visibilidad de los rasgos faciales en condiciones de baja luz, favoreciendo el funcionamiento del sistema en escenarios menos controlados.

En conjunto, estos ajustes permitieron alcanzar un mejor equilibrio entre rendimiento y estabilidad. El sistema logró operar de manera continua, con menor riesgo de interrupciones, lo cual fue fundamental para su posterior evaluación en pruebas de campo, donde las condiciones de operación son más demandantes.

3.3 Resultados en pruebas de campo

Posteriormente, la propuesta desarrollada fue evaluada en escenarios no controlados, con el propósito de simular condiciones reales de operación. En esta etapa se excluyó el caso C5, debido a que representa una condición en la que no es

posible realizar detección basada en características visuales. Para cada combinación de ambiente y caso, se realizaron 3 repeticiones experimentales, con el objetivo de reducir la variabilidad inherente a los entornos no controlados, como cambios de iluminación, ruido en la escena o ligeras variaciones en la posición del sujeto, y así obtener resultados más representativos y confiables del desempeño del sistema. En total, se llevaron a cabo 60 pruebas.

En la Figura 5 se muestran los resultados obtenidos durante las pruebas de campo.

Figura 5

Resultados de detección del sistema híbrido en pruebas de campo.

Ambiente \ Caso	C1	C2	C3	C4	C6
E1 (Oscuro)	V	V	F	F	V
E1 (Repetición)	V	V	V	F	F
E1 (Repetición 2)	V	V	V	F	F
E2 (Muy iluminado)	V	V	V	V	V
E2 (Repetición)	V	V	V	V	V
E2 (Repetición 2)	V	V	V	V	V
E3 (Iluminación media)	V	V	V	F	V
E3 (Repetición)	V	V	V	V	V
E3 (Repetición 2)	V	V	V	V	V
E4 (Media + objetos)	V	V	F	V	F
E4 (Repetición)	V	V	V	V	V
E4 (Repetición 2)	V	V	V	V	V

Los resultados muestran que el método propuesto alcanzó un total de 51 detecciones correctas y 9 incorrectas, lo que corresponde a una exactitud aproximada del 85 %.

Se observa un mejor desempeño en los ambientes con iluminación adecuada, particularmente en aquellos con iluminación media y alta, donde la mayoría de los casos fueron detectados correctamente. En condiciones de baja iluminación, la propuesta mantuvo un rendimiento aceptable, aunque presentó algunos fallos en situaciones con menor visibilidad.

Por otro lado, en escenarios con presencia de objetos en el entorno, se registraron algunas fallas adicionales, especialmente en casos donde el rostro presentaba una orientación parcial o menor contraste, lo que evidencia la influencia de las condiciones ambientales sobre el desempeño del detector.

3.4 Exactitud por ambiente y caso de prueba

Con el objetivo de analizar con mayor detalle el comportamiento del sistema, se calculó la exactitud de detección para cada combinación de ambiente y caso de prueba, a partir de los resultados obtenidos previamente en las Tablas 1 y 2.

Los valores porcentuales se obtuvieron considerando el número de detecciones correctas (V) respecto al total de repeticiones realizadas en cada escenario específico. Es decir, para cada combinación ambiente-caso, la exactitud se calculó dividiendo el número de detecciones correctas entre el total de pruebas

ejecutadas bajo esas mismas condiciones, aplicando la expresión mostrada en (1).

De esta manera, la tabla de exactitud no introduce datos nuevos, sino que representa una transformación directa de los resultados experimentales registrados de forma binaria (V/F), permitiendo una interpretación más clara del desempeño del sistema.

En la Figura 6 se muestra la exactitud del sistema desglosada por ambiente y caso de prueba, expresada en porcentaje. Asimismo, se observa que los ambientes con condiciones de iluminación favorables (E2 y E3) presentan una exactitud del 100 % en todos los casos evaluados, lo que confirma que el sistema opera de manera óptima cuando las características faciales son claramente visibles y que en el ambiente E1 (oscuro) el sistema mantiene un desempeño alto, aunque con ligeras disminuciones en los casos C3 y C4, lo cual puede atribuirse a la reducción de contraste en los rasgos faciales. Por su parte, el ambiente E4, que incluye la presencia de objetos en el entorno, muestra una mayor variabilidad en los resultados, especialmente en el caso C4, donde la combinación de oclusión y elementos del fondo afecta la detección.

Figura 6

Efectividad del sistema por ambiente y caso de prueba.

Ambiente / Caso	C1	C2	C3	C4	C6	Exactitud promedio (%)
E1 (oscuro)	100%	100%	67%	67%	100%	86.80%
E2 (muy iluminado)	100%	100%	100%	100%	100%	100%
E3 (iluminación media)	100%	100%	100%	100%	100%	100%
E4 (media con objetos)	100%	100%	67%	33%	100%	80%

3.5 Análisis comparativo del desempeño del sistema

La comparación entre los resultados obtenidos en entornos controlados y en pruebas de campo permite analizar el comportamiento general del sistema híbrido después de la etapa de ajustes y optimización. En la Figura 7 se muestra una comparación general de los resultados obtenidos durante ambas etapas experimentales.

Figura 7

Comparación general de resultados obtenidos en entornos controlados y pruebas de campo.

Sistema	Total de pruebas	V sin C5	F sin C5	Exactitud sin C5
Sistema Híbrido (Haar + LBP)	48	36	4	90%
Sistema Final (Haar + LBP, campo)	60	51	9	85%

Como se observa en la Figura 7, el sistema híbrido obtuvo una efectividad del 90 % en las pruebas realizadas sin considerar el caso C5, mientras que en las pruebas de campo se alcanzó una efectividad del 85 %. Aunque ambos resultados muestran un desempeño elevado, la ligera disminución observada en las

pruebas de campo evidencia la influencia de factores propios de los entornos no controlados, como variaciones de iluminación, presencia de objetos, obstrucciones y cambios en la orientación del rostro.

Entre los ajustes más relevantes implementados en el sistema se encuentran la reducción de la resolución de procesamiento, la limitación de cuadros por segundo y la calibración del umbral de brillo. Estas modificaciones permitieron disminuir la carga computacional y mantener una ejecución más estable, favoreciendo una detección más consistente en tiempo real.

También, los resultados muestran que el sistema mantiene un comportamiento estable ante variaciones moderadas de iluminación, especialmente en ambientes con iluminación media y alta. Esto representa una característica importante para aplicaciones de búsqueda y rescate, donde las condiciones visuales pueden cambiar constantemente.

3.6 Limitaciones del sistema

A pesar de los resultados obtenidos, es importante señalar algunas limitaciones del sistema propuesto. En particular, se observó que el desempeño disminuye cuando la visibilidad del rostro es reducida, ya sea por condiciones de iluminación muy bajas o por la presencia de oclusiones parciales.

Uno de los casos más representativos es el caso C5, en el cual el rostro se encuentra completamente cubierto. En este escenario, el sistema no cuenta con información visual suficiente para identificar rasgos faciales como ojos, nariz o contornos, los cuales son necesarios tanto para el detector Viola-Jones como para el descriptor LBP. Por esta razón, durante las pruebas de laboratorio este caso no presentó detecciones correctas.

Con base en este comportamiento, se decidió omitir el caso C5 en las pruebas de campo. Esta decisión se tomó debido a que se trata de una condición en la que el sistema no puede operar, ya que no existe información facial disponible para ser procesada. Es decir, no se trata de un problema de desempeño del modelo, sino de una limitación propia del enfoque utilizado.

Además, incluir este caso dentro del análisis habría afectado la interpretación de los resultados, al incorporar fallos que no dependen del funcionamiento del sistema, sino de la ausencia total de información visual. Por ello, su exclusión permite evaluar de forma más precisa el comportamiento del sistema en escenarios donde existe al menos una parte del rostro visible.

De esta manera, el análisis se centra en condiciones más cercanas a situaciones reales de aplicación, donde la visibilidad del rostro suele ser parcial y no completamente nula.

IV. CONCLUSIONES

El presente trabajo permitió desarrollar e implementar un sistema de visión artificial orientado a la detección de rostros humanos en escenarios simulados de derrumbe, empleando un

enfoque híbrido basado en los métodos Viola-Jones y Local Binary Patterns (LBP). A partir de los resultados obtenidos, se puede afirmar que el sistema cumple con el objetivo principal de realizar detección facial en tiempo real bajo condiciones variables de iluminación, presencia de objetos y visibilidad parcial del rostro.

Uno de los aspectos más relevantes del estudio fue la validación del enfoque híbrido propuesto. La combinación de Viola-Jones para la detección inicial y LBP como etapa de verificación permitió mejorar la confiabilidad del sistema, reduciendo detecciones incorrectas y manteniendo un equilibrio adecuado entre precisión y velocidad. Este comportamiento resulta especialmente importante en aplicaciones de búsqueda y rescate, donde el tiempo de respuesta y la eficiencia computacional son factores críticos.

Los resultados experimentales mostraron que el sistema alcanza una exactitud aproximada del 85 % en pruebas realizadas en entornos no controlados. Este desempeño confirma que el sistema es capaz de adaptarse a condiciones más cercanas a un escenario real, manteniendo una detección estable en casos donde el rostro es visible total o parcialmente. Asimismo, se observó una alta consistencia en los resultados al repetir las pruebas bajo las mismas condiciones, lo que respalda la confiabilidad del sistema desarrollado.

No obstante, también se identificaron limitaciones importantes. El sistema presenta dificultades en escenarios donde el rostro se encuentra completamente cubierto o cuando las condiciones de iluminación son extremadamente deficientes, lo cual es coherente con la naturaleza de los métodos utilizados, que dependen de la presencia de características faciales visibles. Estas limitaciones representan un área clara de mejora para futuros desarrollos.

En términos de implementación, el sistema demostró ser una solución viable de bajo costo computacional, capaz de operar en tiempo real mediante el uso de herramientas accesibles como Python y OpenCV. Esto refuerza su potencial para ser integrado en plataformas embebidas o sistemas robóticos destinados a tareas de exploración en entornos de riesgo.

Como trabajo futuro, se propone la integración del sistema en un robot terrestre que permita evaluar su desempeño en condiciones dinámicas, considerando factores como movimiento, vibraciones y cambios constantes en el entorno. Asimismo, resulta pertinente explorar la incorporación de técnicas de aprendizaje profundo optimizadas, así como la ampliación de la base de datos de entrenamiento con escenarios más complejos que incluyan polvo, humo y oclusiones más severas. Por último, el uso de métricas de evaluación más detalladas permitirá analizar con mayor precisión el comportamiento del sistema y orientar mejoras enfocadas en la reducción de falsos negativos, los cuales son críticos en aplicaciones de rescate.

En conjunto, los resultados obtenidos demuestran que el sistema propuesto constituye una alternativa funcional y prometedora

para apoyar tareas de búsqueda y rescate, sirviendo como base para desarrollos futuros orientados a mejorar la detección de personas en entornos complejos.

V. AGRADECIMIENTOS

Agradezco al Dr. Héctor A. Moreno Casillas por la orientación y el acompañamiento brindado durante el desarrollo de la tesis de la cual se deriva este artículo, ya que su apoyo fue fundamental en la construcción de las bases de este trabajo.

Asimismo, agradezco al Instituto Tecnológico de La Laguna por las facilidades y el entorno académico proporcionado para la realización de este proyecto.

De igual manera, expreso mi agradecimiento a los docentes de la línea de Instrumentación Electrónica del Departamento de Posgrado de la Maestría en Ciencias en Ingeniería Eléctrica, por los conocimientos compartidos y la formación recibida, los cuales contribuyeron directamente al desarrollo de esta investigación.

También agradezco a las personas que participaron en las pruebas experimentales, ya que su colaboración permitió evaluar el sistema en distintos escenarios y obtener resultados más completos.

Extiendo mi agradecimiento a quienes, de alguna manera, brindaron apoyo durante el desarrollo del proyecto, ya sea con sugerencias, apoyo técnico o disponibilidad para la realización de las pruebas.

VI. REFERENCIAS

- Adeshina, S. O., Ibrahim, H., Teoh, S. S., & Hoo, S. C. (2021). Custom face classification model for classroom using Haar-like and LBP features with their performance comparisons. *Electronics*, 10(2), 102. <https://doi.org/10.3390/electronics10020102>
- Ahonen, T., Hadid, A., & Pietikäinen, M. (2006). Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28 (12), 2037–2041. <https://doi.org/10.1109/TPAMI.2006.244>
- AlMousawi, D. K. M., Hassan, A. A., & Hayawi, D. M. J. (2020). Hybrid method for face description using LBP and HOG. *Journal of Education and Pure Science*, 10(1), 73–79.
- Ambrogio, E. Á. J. (2020). Reconocimiento de objetos a través de la metodología Haar Cascades. *Tendencias en Ingeniería*, 8(1).
- Centro Nacional de Prevención de Desastres. (19 de septiembre de 2019). *Los sismos del 19 de septiembre*. Gobierno de México. <https://www.gob.mx/cenapred/articulos/los-sismos-del-19-septiembre?idiom=es>
- Erdelj, M., Król, M., & Natalizio, E. (2017). Wireless sensor networks and multi-UAV systems for natural disaster management. *Computer Networks*, 124, 72–86. <https://doi.org/10.1016/j.comnet.2017.05.021>
- García-Olalla, O., & Alegre, E. (s. f.). Descripción de textura en imágenes utilizando Local Binary Pattern (LBP). En *Conceptos y métodos en visión por computador* (Cap. 7). Universidad de León. <https://buleria.unileon.es/bitstream/handle/10612/11063/cap%20%20ConceptosyMetodosenVxC.pdf>
- Gerónimo, D., López, A. M., Sappa, A. D., & Graf, T. (2010). Survey of pedestrian detection for advanced driver assistance systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(7), 1239–1258. <https://doi.org/10.1109/TPAMI.2009.122>
- Jadeja, R., Trivedi, T., & Surve, J. (2024). Survivor detection approach for post-earthquake search and rescue missions based on deep learning inspired algorithms. *Scientific Reports*, 14(1), 25047. <https://doi.org/10.1038/s41598-024-75156-z>
- Liu, Z., Zhang, Y., & Liu, J. (2020). Vision-based human detection for search and rescue robots in disaster environments. *IEEE Access*, 8, 146600–146611. <https://doi.org/10.1109/ACCESS.2020.3014858>
- Moreno, R. J. (2014). Robotic explorer to search people through face detection. En *International Congress of Engineering Mechatronics and Automation (CIIMA)*. IEEE. <https://doi.org/10.1109/CIIMA.2014.6983429>
- Murphy, R. R., Tadokoro, S., Kleiner, A., Nardi, D., Jacoff, A., Messina, E., & Erkmén, A. M. (2008). Disaster robotics. En B. Siciliano & O. Khatib (Eds.), *Springer handbook of robotics* (pp. 1151–1173). Springer. https://doi.org/10.1007/978-3-540-30301-5_51
- Ojala, T., Pietikäinen, M., & Mäenpää, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7), 971–987. <https://doi.org/10.1109/TPAMI.2002.1017623>
- Restas, A. (2015). Drone applications for supporting disaster management. *World Journal of Engineering and Technology*, 3, 316–321. <https://doi.org/10.4236/wjet.2015.33C047>
- Sutikno, S., Sugiharto, A., Kusumaningrum, R., & Wibawa, H. A. (2025). Combination of HAAR, HOG, and LBP descriptors for enhanced classification of moving objects. *Engineering, Technology & Applied Science Research*, 15(3).
- Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. En *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)* (Vol. 1, pp. 511–518). IEEE. <https://doi.org/10.1109/CVPR.2001.990517>
- Xu, Y., Yu, G., Wang, Y., Wu, X., & Ma, Y. (2016). A hybrid vehicle detection method based on Viola–Jones and HOG + SVM from UAV images. *Sensors*, 16(8), 1325. <https://doi.org/10.3390/s16081325>
- Zhao, Z.-A. (2024). Enhancing human detection in occlusion-heavy disaster scenarios: A visibility-enhanced DINO model. *Smart Cities*, 8(1), 12. <https://doi.org/10.3390/smartcities8010012>

No.	Rol	Descripción
1	Conceptualización [Conceptualization]	Hasem Azareel Gallegos Rodríguez
2	Curación de datos [Data curation]	Hasem Azareel Gallegos Rodríguez
3	Análisis formal [Formal analysis]	Hasem Azareel Gallegos Rodríguez
4	Obtención de fondos [Funding acquisition]	Hasem Azareel Gallegos Rodríguez
5	Investigación [Investigation]	Hasem Azareel Gallegos Rodríguez
6	Metodología [Methodology]	Hasem Azareel Gallegos Rodríguez
7	Administración del proyecto [Project administration]	Hasem Azareel Gallegos Rodríguez
8	Recursos [Resources]	Hasem Azareel Gallegos Rodríguez
9	Software [Software]	Hasem Azareel Gallegos Rodríguez
10	Supervisión [Supervision]	Hasem Azareel Gallegos Rodríguez
11	Validación [Validation]	Hasem Azareel Gallegos Rodríguez
12	Visualización [Visualization]	Hasem Azareel Gallegos Rodríguez
13	Redacción - borrador original [Writing – original draft]	Hasem Azareel Gallegos Rodríguez
14	Redacción - revisión y edición [Writing – review & editing]	Hasem Azareel Gallegos Rodríguez