

---

# Optimización de Sistemas de Reconocimiento Facial Para Entornos Escolares : Un Estudio de Modelos CNN y HOG (Noviembre 2024)

**Angel Daniel Grimaldo Ocampo**

Universidad Autónoma de Baja California

Facultad de Ingeniería Arquitectura y Diseño  
Ingeniería en Software y Tecnologías Emergentes

e-mail : angel.grimaldo@uabc.edu.mx

## ABSTRACT

La tecnología de reconocimiento facial presenta una solución eficiente y no intrusiva para la automatización del pase de lista en entornos educativos, abordando las limitaciones de los métodos manuales tradicionales. Este estudio evalúa el desempeño de dos modelos ampliamente utilizados en reconocimiento facial: Redes Neuronales Convolucionales (CNN) y Histogramas de Gradientes Orientados (HOG), en un entorno controlado optimizado con iluminación favorable y mejoras clave en el sistema. Estas mejoras incluyen validación doble, reducción de la tolerancia en la detección y entrenamiento limitado a imágenes específicas de un grupo.

Los resultados obtenidos muestran un aumento significativo en las métricas de desempeño, alcanzando un 96.42% de exactitud y 95.07% de precisión para CNN y 95.84% de exactitud y 94.07% de precisión para HOG, lo que posiciona a este último como una alternativa viable en escenarios con recursos limitados. Este estudio destaca las implicaciones prácticas de implementar sistemas de reconocimiento facial en las escuelas, demostrando su capacidad para optimizar los procesos de asistencia mientras garantiza accesibilidad mediante tecnologías más económicas. Se discuten las limitaciones relacionadas con entornos controlados y configuraciones estáticas, sentando las bases para investigaciones futuras orientadas a aplicaciones más amplias.

## INDEX TERMS

algoritmos, cnn, dataset, exactitud, precisión, face recognition, hog, modelos, opencv, python, redes neuronales, visión computacional, entornos controlados, tecnología educativa.

## I. INTRODUCCIÓN

En el ámbito educativo, la automatización de procesos como el pase de lista se presenta como una solución innovadora para optimizar la gestión del tiempo y mejorar la precisión en el registro de asistencia. El reconocimiento facial, como herramienta tecnológica, ha ganado relevancia debido a su capacidad para identificar individuos de manera eficiente y no invasiva. Sin embargo, su implementación en entornos educativos plantea desafíos técnicos, entre ellos, la selección del modelo de detección y reconocimiento facial

más adecuado para equilibrar precisión, exactitud y requerimientos computacionales.

En investigaciones previas se identificó que los modelos basados en Redes Neuronales Convolucionales (CNN) superaban a los Histogramas de Gradientes Orientados (HOG) en exactitud bajo condiciones de iluminación variables, mientras que las diferencias en precisión eran mínimas. A partir de estos hallazgos, la presente investigación modifica el enfoque para evaluar el rendimiento de ambos modelos en un entorno controlado que incorpora mejoras significativas al sistema, como la

---

doble validación, la reducción del rango de tolerancia en detección facial y el entrenamiento limitado a imágenes específicas de un grupo.

En el contexto actual, los métodos manuales de pase de lista en entornos escolares no solo consumen tiempo valioso, sino que también son propensos a errores humanos. A pesar de los avances tecnológicos, no existe una solución ampliamente adoptada que aproveche el modelo HOG, conocido por su eficiencia computacional, en sistemas de reconocimiento facial para este propósito. Esto deja una brecha importante en el desarrollo de tecnologías accesibles y de alto rendimiento que puedan ser implementadas en instituciones educativas con recursos limitados.

## Objetivos

### Objetivo General:

Evaluar el desempeño de los modelos HOG y CNN en un sistema automatizado de pase de lista en entornos educativos, tras la implementación de mejoras específicas que optimizan la detección facial.

### Objetivos Específicos:

1. Validar el impacto de las mejoras implementadas en el sistema, como la doble validación y la reducción de tolerancia en detección facial.
2. Analizar la viabilidad de HOG como alternativa a CNN en escenarios con recursos computacionales limitados.
3. Comparar la exactitud y precisión de los modelos HOG y CNN bajo condiciones de luz controladas.

### Impacto Organizacional e Implicaciones Prácticas:

La implementación de un sistema de reconocimiento facial automatizado en instituciones educativas puede optimizar significativamente el tiempo dedicado al registro de asistencia, permitiendo a los docentes concentrarse en actividades pedagógicas. Además, al garantizar una mayor precisión en los registros, se eliminan problemas asociados a la asistencia incorrecta, facilitando una gestión más eficiente y transparente.

### Impacto Tecnológico:

El desarrollo de un sistema que integre el modelo HOG para detección facial, optimizado con mejoras como la doble validación, aporta una solución innovadora y accesible, destacando por su bajo consumo de recursos computacionales. Esto amplía las posibilidades de implementación en instituciones con limitaciones de infraestructura tecnológica, fomentando la democratización del acceso a tecnologías avanzadas.

### Alcance y Limitaciones:

El sistema desarrollado está diseñado para entornos controlados con condiciones óptimas de iluminación y requiere que los alumnos se posicionen frente a la cámara en un rango limitado. Esto asegura la máxima efectividad del reconocimiento facial, pero limita su aplicabilidad en entornos más dinámicos o con condiciones de luz no ideales. Además, aunque los resultados demuestran la viabilidad de HOG como alternativa a CNN, futuras investigaciones deben explorar su rendimiento en bases de datos más amplias y diversos contextos educativos.

Este trabajo no solo aborda un problema práctico en el sector educativo, sino que también contribuye al avance del campo de la visión computacional, evaluando y optimizando tecnologías existentes para maximizar su utilidad en escenarios del mundo real.

## II. TRABAJO RELACIONADO

### Sistemas de Reconocimiento Facial Aplicados en Entornos Escolares

El uso de tecnologías de reconocimiento facial en entornos educativos ha sido explorado previamente, con énfasis en modelos como CNN y otros algoritmos clásicos. Un estudio basado en CNN obtuvo un 95% de precisión al implementar un sistema de asistencia automatizado, destacando su alta exactitud en la extracción de características faciales complejas. Sin embargo, el costo computacional y los recursos necesarios limitan su aplicabilidad en contextos educativos con hardware básico [1].

Por otro lado, se han utilizado algoritmos como LBPH y Viola-Jones para diseñar sistemas más accesibles. Un sistema basado en LBPH logró procesar datos de manera eficiente y almacenar información en hojas de cálculo, pero carece de capacidades avanzadas de integración y generación de reportes en tiempo real [2]. Asimismo, el algoritmo Viola-Jones alcanzó precisiones de entre 90% y 94%, siendo una solución viable para detección facial en tiempo real, aunque no aborda problemas de interoperabilidad en plataformas web [3].

Sin embargo, no se encontraron investigaciones específicas sobre el uso del modelo HOG en sistemas escolares, lo que resalta la novedad de este proyecto al explorar su aplicabilidad en este ámbito, aprovechando su bajo consumo de recursos computacionales en comparación con CNN.

### Investigaciones Sobre Mejoras a los Algoritmos CNN y HOG

En el contexto de optimización, un estudio integró CNN con LSTM, alcanzando un 99% de precisión. Aunque los resultados fueron impresionantes, la implementación requirió recursos avanzados y tiempos de ejecución prolongados, factores poco ideales para entornos escolares [4]. Otro trabajo mejoró la detección facial en escenarios complejos con MTCNN, comparándolo con HOG y Haar Cascade, subrayando la eficiencia de HOG en entornos controlados, aunque con limitaciones frente a modelos más avanzados en condiciones cambiantes [5].

No se encontraron estudios que implementen específicamente las mejoras propuestas en este proyecto, como la doble validación, el ajuste de parámetros y la limitación del entrenamiento a imágenes del grupo objetivo, lo que resalta la contribución original de esta investigación.

### Comparación Entre Modelos CNN y HOG en Reconocimiento Facial

La comparación directa entre CNN y HOG ha sido explorada principalmente en contextos generales, no escolares. En un estudio, se destacó que CNN maneja mejor las variaciones de iluminación y perspectiva, pero HOG ofrece una relación costo-beneficio superior para aplicaciones con recursos limitados [6]. Otro análisis subrayó que HOG, aunque menos preciso en escenarios adversos, es una alternativa viable y eficiente en entornos controlados [7].

En el ámbito educativo, no se identificaron investigaciones que comparen explícitamente HOG y CNN, dejando un vacío que este proyecto busca abordar al evaluar su desempeño en un entorno escolar controlado, implementando un sistema accesible y eficiente.

## III. HERRAMIENTAS PARA EL DESARROLLO DEL SISTEMA

La visión computacional, un área clave de la Inteligencia Artificial, permite a las computadoras analizar y extraer información de imágenes para resolver problemas prácticos. Este proyecto aprovecha herramientas tecnológicas avanzadas para implementar un sistema de pase de lista automatizado basado en reconocimiento facial, empleando un enfoque integrado de desarrollo que combina bibliotecas de Python, un entorno web moderno y una base de datos para la gestión de información.

### Herramientas Utilizadas

#### Librerías de Python

- **OpenCV:** Esta biblioteca es fundamental en el campo de la visión computacional,

proporcionando funciones esenciales para el procesamiento de imágenes, como el filtrado, la detección de bordes, el reconocimiento de características y el seguimiento de objetos. En este sistema, OpenCV se utiliza para capturar imágenes en tiempo real desde la cámara integrada de la computadora y realizar preprocesamiento básico antes de la detección facial.

- **Face Recognition:** Complementa el uso de OpenCV al encargarse específicamente de la detección y comparación de rostros. Esta biblioteca implementa dos modelos principales para la detección facial:

-CNN (Convolutional Neural Network): Un modelo que utiliza redes neuronales profundas para extraer y aprender características complejas, logrando una alta exactitud en la detección facial.

-HOG (Histograms of Oriented Gradients): Un modelo que emplea técnicas más tradicionales de análisis de gradientes para identificar rostros en las imágenes.

### Entorno Web y Gestión de Datos

- **React y Next.js:** Se utilizó React para la creación de una interfaz de usuario interactiva y moderna, y Next.js para gestionar el lado del servidor y facilitar la integración con la base de datos. Estas herramientas permiten la interacción entre los usuarios del sistema (profesores y alumnos) y el módulo de reconocimiento facial, asegurando una experiencia fluida.
- **Base de Datos:** Se desarrolló un sistema de almacenamiento centralizado que permite el intercambio de información entre la interfaz web y el modelo de Python. La base de datos almacena las imágenes de los alumnos en dos versiones (con y sin lentes), los registros de asistencia generados por el sistema y los parámetros configurables para la detección facial.

### Hardware Utilizado

El sistema fue implementado y evaluado en una computadora MacBook Pro con chip M3 Pro, aprovechando su cámara integrada de 1080p y su capacidad computacional para ejecutar modelos de visión por computadora de forma eficiente.

### Enfoque del Proyecto

Las herramientas seleccionadas trabajan de manera integrada para optimizar el desempeño del sistema. La interfaz web facilita la doble validación de los alumnos mediante el uso de dos imágenes (con y sin lentes) y restringe el entrenamiento del modelo a imágenes de un grupo específico. Además, el ajuste en el nivel de tolerancia para la detección facial (de 0.6 a 0.44) se implementó directamente en el modelo gestionado por la librería Face Recognition, logrando una reducción significativa en los falsos positivos y mejorando la confiabilidad del sistema.

Este enfoque tecnológico asegura un entorno controlado y eficiente, eliminando problemas previos como confusiones por cambios en las condiciones de iluminación o por la variación de características en las imágenes capturadas.

#### IV. CNN & HOG

CNN y HOG son técnicas de procesamiento de imágenes utilizadas para la detección y reconocimiento de objetos, incluyendo rostros. Ambas técnicas extraen características distintivas de las imágenes para facilitar la identificación.

##### A. CNN

Las CNN son un tipo de red neuronal profunda especialmente eficaz para tareas de reconocimiento de imágenes. Utilizan capas convolucionales para extraer características de las imágenes de entrada, seguido de capas de pooling para reducir la dimensionalidad y capas completamente conectadas para la clasificación. Las CNN son capaces de aprender características complejas directamente de los datos crudos, lo que las hace muy precisas en tareas de detección y reconocimiento de imágenes.

##### Ventajas

**Alta Precisión:** Las CNN suelen tener una muy buena precisión en la detección y reconocimiento facial debido a su capacidad para aprender características complejas.

**Generalización:** Pueden generalizar bien a nuevas imágenes, incluso si las condiciones de iluminación o las posiciones de las caras varían.

##### Desventajas

**Requerimientos Computacionales:** Las CNN requieren una gran cantidad de recursos computacionales y tiempo para el entrenamiento.

**Tamaño del Modelo:** Los modelos de CNN suelen ser grandes, lo que puede ser un problema para dispositivos con limitaciones de memoria o potencia de procesamiento.

##### Funcionamiento

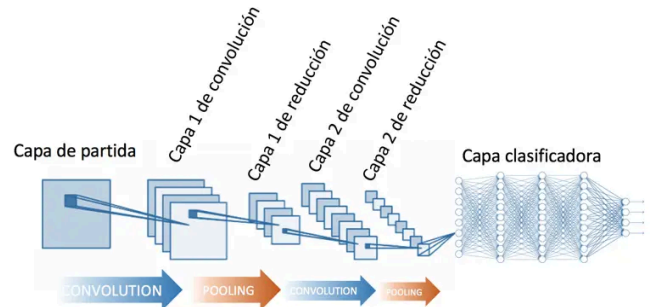


FIGURA 1. Funcionamiento del algoritmo de CNN.

**Capa de partida:** Es la imagen de entrada.

**Capas de convolución:** Aplican filtros para extraer características de la imagen, como bordes y texturas.

**Capas de reducción (Pooling):** Reducen la dimensionalidad y retienen las características más importantes, mejorando la eficiencia.

**Capa clasificadora:** Es una red neuronal totalmente conectada que utiliza las características extraídas para clasificar la imagen en diferentes categorías.

##### B. HOG

HOG es un descriptor de características que cuenta la ocurrencia de gradientes de orientación en localizaciones específicas de una imagen. Este enfoque se basa en la idea de que la apariencia y la forma de los objetos pueden caracterizarse bien por la distribución de los gradientes de intensidad o las direcciones de borde en una imagen. HOG es rápido de calcular y es eficiente en términos de uso de recursos, lo que lo hace adecuado para aplicaciones en tiempo real.

##### Ventajas

**Eficiencia:** HOG es más eficiente en términos de tiempo de cómputo y uso de recursos, lo que lo hace adecuado para aplicaciones en tiempo real.

**Simplicidad:** Es más fácil de implementar y ajustar en comparación con las CNN.

## Desventajas

Menor Precisión: HOG generalmente tiene una precisión menor en comparación con las CNN, en condiciones de variabilidad como cambios de iluminación o poses faciales. Dependencia de la Calidad de Imagen: Su rendimiento puede verse significativamente afectado por la calidad de las imágenes de entrada.

## Funcionamiento

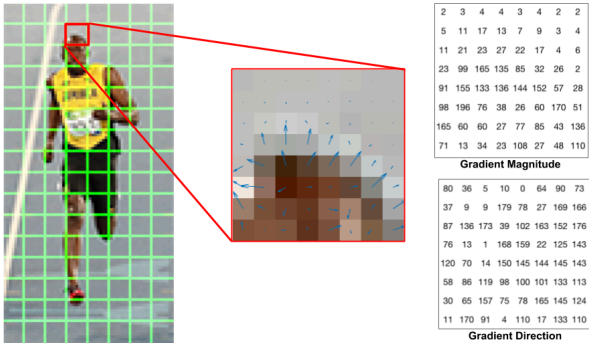


FIGURA 2. Funcionamiento del algoritmo de HOG.

**División de la imagen:** La imagen se divide en pequeñas celdas.

**Cálculo de gradientes:** Se calculan los gradientes de orientación en cada celda.

**Construcción de histogramas:** Se construye un histograma de las orientaciones de los gradientes en cada celda.

**Normalización de bloques:** Se agrupan celdas en bloques y se normalizan para mejorar la invariancia a la iluminación.

**Concatenación de características:** Los histogramas normalizados se concatenan para formar un vector de características que representa la imagen.

## RESUMEN

Ambos algoritmos son capaces de cumplir con la tarea de la detección de rostros y su clasificación, sin embargo, independientemente de su funcionamiento, hay dos puntos esenciales a destacar de ambos algoritmos: su precisión y requerimientos computacionales necesarios para su ejecución. En teoría Hog es rápido, pero se le pueden dificultar la detección de algunas imágenes, mientras que CNN requiere muchos recursos pero su detección es precisa.

## V. METODOLOGÍA

El desarrollo del proyecto se llevó a cabo utilizando la metodología ágil Scrum, adaptada para un entorno de trabajo individual. Este enfoque permitió estructurar el trabajo en ciclos iterativos llamados sprints, priorizando tareas específicas en cada iteración para asegurar un progreso constante y enfocado hacia los objetivos del sistema. A continuación, se describen las fases principales del desarrollo:

### Rediseño del Sistema de Reconocimiento Facial

El primer sprint se centró en reestructurar el sistema de Python utilizando librerías como OpenCV y FaceRecognition. En esta etapa, se revisaron y seleccionaron los modelos HOG (Histograms of Oriented Gradients) y CNN (Convolutional Neural Networks), evaluando sus fortalezas y limitaciones. Se realizaron ajustes en los parámetros de detección facial, como la reducción de la tolerancia de detección al 44%, y se limitó el entrenamiento a un conjunto específico de imágenes para maximizar la exactitud en el entorno controlado.

### Optimización y Mejora del Algoritmo

En los sprints subsiguientes, se introdujeron mejoras adicionales en el sistema, como la implementación de una validación doble para evaluar imágenes con y sin lentes, incrementando así la robustez ante variaciones en las características de los usuarios. Este proceso fue iterativo, permitiendo realizar pruebas continuas con métricas clave (VP, FP, VN, FN) para refinar los ajustes del sistema hasta alcanzar los resultados deseados.

### Desarrollo del Sistema Web y CRUDs

Paralelamente, se desarrolló la plataforma web utilizando React y Next.js. Se inició con la creación de los módulos de gestión de datos (CRUDs) para entidades clave como alumnos, grupos, eventos, profesores y administradores. Posteriormente, se diseñaron interfaces personalizadas para cada rol, como páginas de gestión de asistencia para

profesores y vistas de reportes para administradores. Este trabajo se realizó de manera incremental, mejorando la funcionalidad y el diseño con cada iteración.

### Integración y Pruebas del Sistema

En la fase de integración, se conectaron el sistema de Python y la plataforma web mediante una base de datos centralizada que permitió la comunicación eficiente entre ambos. Las pruebas se realizaron asegurando la funcionalidad correcta de los flujos de datos y la interacción entre los roles de usuario. Además, el sistema fue probado en un entorno controlado con iluminación favorable, garantizando su correcto funcionamiento bajo las condiciones previstas.

### Evaluación y Ajustes Finales

En el sprint final, se evaluaron el rendimiento y las métricas de los modelos HOG y CNN utilizando un dataset de prueba compuesto por imágenes de rostros conocidos y desconocidos para los algoritmos entrenados. Se calcularon las métricas de exactitud y precisión, obteniendo resultados superiores a los de investigaciones previas gracias a las optimizaciones realizadas. El modelo HOG alcanzó una exactitud del 95.84% y una precisión del 94.07%, acercándose significativamente al rendimiento de CNN (96.42% y 95.07% en exactitud y precisión).

Esta metodología permitió realizar ajustes iterativos, realizar pruebas continuas y garantizar que tanto el sistema de reconocimiento facial como la plataforma web cumplieran con los objetivos planteados.

Al igual que con la investigación pasada, se creó otro dataset con 30 registros los cuales comparan las métricas de ambos algoritmos.

## VI. RECOLECCIÓN DE DATOS

Para determinar la exactitud de cada modelo en cada prueba, se probaron ambos modelos con imágenes en donde se mostraban personas en diversas profundidades, orientaciones de perfil del rostro y cantidad de luz en la imagen. El objetivo es evaluar la cantidad de Falsos Positivos, Falsos Negativos, Verdaderos Positivos y Verdaderos Negativos para poder calcular la exactitud y precisión en cada imagen.

Las fórmulas aplicadas para el cálculo de la precisión y exactitud son las siguientes:

$$exactitud = \frac{VP + VN}{VP + FP + FN + VN}$$

FIGURA 3. Fórmula para calcular la exactitud.

$$precision = \frac{VP}{VP + FP}$$

FIGURA 4. Fórmula para calcular la precisión.

La forma en la que se recolectan los datos es mediante el uso de la tabla de asistencia que se genera con la lista de asistencia del programa de Python.

Grupo 23

Lista de Asistencia

ESTUDIANTE	ESTADO
Angel Gimelato	Presente
Gabriel Lizola	Ausente
Gonzalo Espinoza	Ausente
José Díaz	Presente
Juan Luis	Ausente
Carlos Valdejo	Ausente
Hector Zalarán	Ausente
Martin Ortiz	Presente
Sebastian Jimenez	Ausente
Noelia Orozco	Ausente
Estefanía Tello	Ausente
Francisco Cabrales	Ausente
Isaac Gonzalez	Ausente

FIGURA 5. Tabla de Excel generada con las personas detectadas, 1 representa asistencia y 0 inasistencia.

Una vez que se obtiene esta lista, se procede a determinar los siguientes datos:

-Verdaderos Positivos (VP) : Personas en la imagen que fueron detectadas correctamente.

-Verdaderos Negativos (VN) : Personas en la lista de asistencia que no fueron marcadas como detectadas y que no aparecían en la imagen.

-Falsos Positivos (FP) : Personas que se marcaron como detectadas en la lista pero realmente no estaban en la imagen.

-Falsos Negativos (FN) : Personas que aparecían en la imagen pero no fueron detectados en la lista.

Se hicieron pruebas para recolectar estos datos con 30 imágenes en las cuales se encontraban los rostros de algunas de las personas con las que fueron entrenados los modelos.

La mayoría de las imágenes con las que se le entrenó al modelo fueron capturadas por un iPhone 15 Pro Max con la cámara Gran angular de 48 MP: 24 mm, apertura de  $f/1.78$  y la cámara Teleobjetivo de 2x de 12 MP (por medio del sensor quad pixel): 48 mm, apertura de  $f/1.78$ .

El procesamiento de ambos modelos (CNN y HOG) fue hecho con una MacBook Air 2023 con arquitectura ARM y con el chip Apple Silicon M3 Pro.

## VII. RESULTADOS

Tras la implementación de las mejoras propuestas en el sistema, se obtuvieron resultados significativamente superiores en comparación con los alcanzados en investigaciones anteriores. En esta nueva evaluación, las métricas de exactitud y precisión mostraron un incremento notable para ambos modelos, lo que valida la efectividad de las estrategias implementadas, como la doble validación, el ajuste en los parámetros de detección y la reducción de la tolerancia.

### Resultados de Evaluación:

- **Modelo CNN:**
  - **Exactitud:** 96.42%
  - **Precisión:** 95.07%
- **Modelo HOG:**
  - **Exactitud:** 95.84%
  - **Precisión:** 94.07%

En comparación con la investigación previa, los resultados reflejan una mejora significativa. En los estudios iniciales, el modelo CNN alcanzó una exactitud del 91% y una precisión del 72%, mientras que HOG obtuvo una exactitud del 89% y una precisión del 70%. La optimización del sistema permitió que ambos modelos alcanzaran niveles de desempeño casi idénticos, con una ligera ventaja para CNN en términos de precisión.

Estos resultados destacan la importancia de las mejoras implementadas en el diseño del sistema, incluyendo un entorno controlado, restricciones de posicionamiento de los usuarios y un ajuste riguroso de los parámetros de detección. Además, las pruebas realizadas con la MacBook M3 Pro y su cámara integrada de 1080p proporcionaron una base sólida para garantizar la consistencia de los datos obtenidos.

Se utilizó el test de Wilcoxon para evaluar la diferencia significativa en las métricas de precisión y exactitud del modelo CNN y el modelo HOG antes y después de los ajustes implementados. El uso de este test no paramétrico se justificó debido a que los datos de las métricas no siguieron una distribución normal, como lo confirmaron los resultados del test de Shapiro-Wilk ( $p$ -valores  $< 0.05$  para todos los casos).

Se creó un dataset con los datos comprendidos por los VP, VN, FN y FP correspondientes a cada imagen y modelo. A cada imagen se le evaluó su precisión y exactitud generada por cada modelo y posteriormente se promedió para determinar la exactitud y precisión promedio ambos modelos.

*Estos fueron los resultados :*

prueba	cnrVP	cnrVN	cnrFP	cnrFN	hogVP	hogVN	hogFP	hogFN	exactitudCnn	exactitudHog	precisionCnn	precisionHog
1	9	9	1	1	7	9	2	2	90	80	90	77.777777
2	9	11	0	0	9	11	0	0	100	100	100	100
3	7	13	0	2	7	13	0	1	90.909091	95.238095	100	100
4	0	36	0	0	0	36	0	0	100	100	100	100
5	0	40	0	0	0	40	0	0	100	100	100	100
6	8	10	0	2	8	10	0	2	90	90	100	100
7	5	15	0	0	5	15	0	0	100	100	100	100
8	5	14	1	0	5	14	1	0	95	95	83.333333	83.333333
9	3	15	0	2	3	15	0	2	90	90	100	100
10	5	15	0	0	5	15	0	0	100	100	100	100
11	5	15	0	0	4	15	0	1	100	95	100	100
12	1	59	1	0	1	59	1	0	98.360657	98.360657	90	90
13	8	10	0	2	7	10	0	3	90	85	100	100
14	5	15	0	0	5	15	0	0	100	100	100	100
15	4	15	0	1	5	15	0	0	95	100	100	100
16	5	15	0	0	5	15	0	0	100	100	100	100
17	1	59	1	0	1	58	2	0	98.360657	96.721314	90	83.333333
18	4	15	0	1	5	15	0	0	95	100	100	100
19	5	15	0	0	5	15	0	0	100	100	100	100
20	8	10	0	2	7	10	0	3	90	85	100	100
21	5	15	0	0	5	15	0	0	100	100	100	100
22	5	15	0	0	5	15	0	0	100	100	100	100
23	9	9	1	1	8	9	1	2	90	85	90	88.888889
24	1	60	0	0	1	60	0	0	100	100	100	100
25	5	15	0	0	5	15	0	0	100	100	100	100
26	10	10	0	0	9	10	0	1	100	95	100	100
27	4	15	0	1	5	15	0	0	95	100	100	100
28	5	15	0	0	5	15	0	0	100	100	100	100
29	5	15	0	0	5	15	0	0	100	100	100	100
30	8	9	1	2	8	9	1	2	85	85	88.888889	88.888889
	154	584	8	17	150	583	8	19	96.421034	95.8400208	95.07407407	94.07407407

FIGURA 6. Tabla de resultados obtenidos por cada modelo.

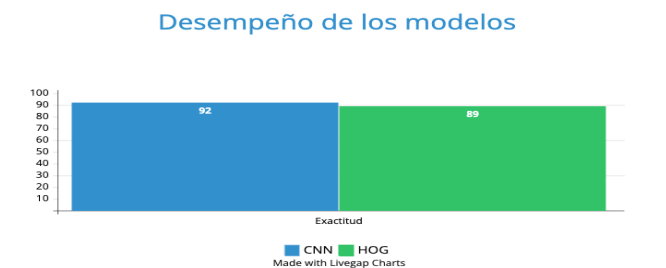


FIGURA 7. Gráfica con el porcentaje de exactitud de ambos modelos sin las mejoras en el sistema anteriormente mencionadas.

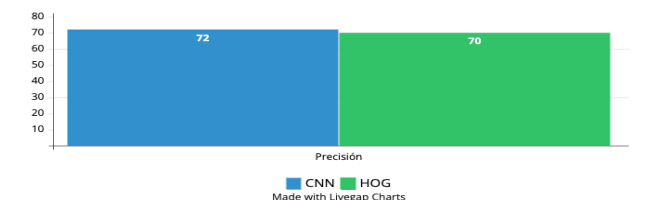


FIGURA 8. Gráfica con el porcentaje de precisión de ambos modelos sin las mejoras en el sistema.

Debe a notar que CNN en ambas gráficas es ligeramente superior a HOG .

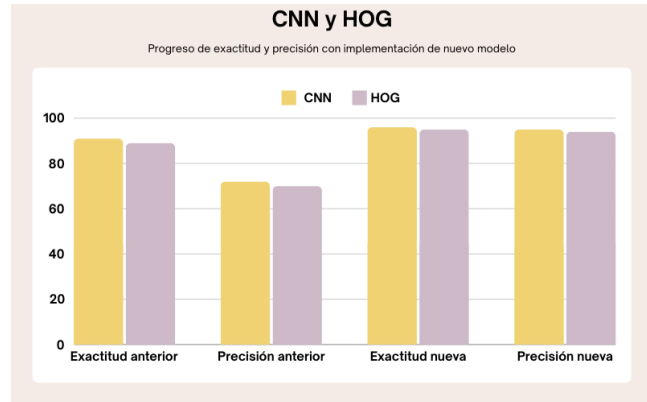


FIGURA 9. Gráficas con el porcentaje de precisión y exactitud de ambos modelos con las mejoras en el sistema.

## Resultados obtenidos

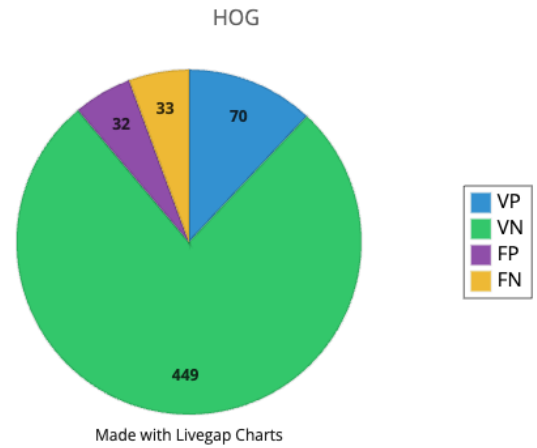


FIGURA 11. Gráfica de pastel de los resultados obtenidos por el modelo HOG sin las mejoras en el sistema.

## Resultados obtenidos

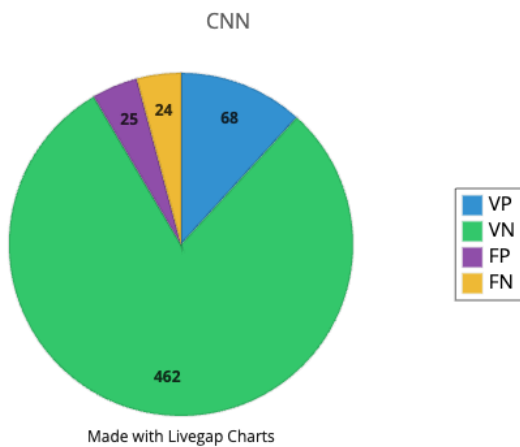


FIGURA 10. Gráfica de pastel de los resultados obtenidos por el modelo CNN sin las mejoras en el sistema.

## CNN

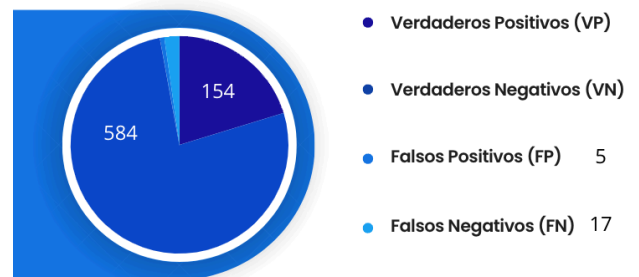


FIGURA 11. Gráfica de pastel de los resultados obtenidos por el modelo CNN con las mejoras en el sistema.

## HOG

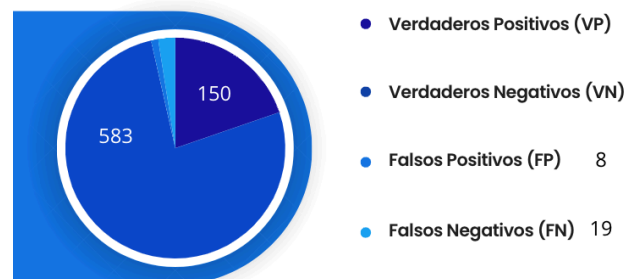


FIGURA 12. Gráfica de pastel de los resultados obtenidos por el modelo HOG con las mejoras en el sistema.

```
> wilcox.test(before, after, paired = TRUE)

Wilcoxon signed rank test with continuity correction

data: before and after
V = 36.5, p-value = 0.003636
alternative hypothesis: true location shift is not equal to 0
```

**FIGURA 12.** Prueba de Wilcoxon para determinar diferencia significativa en la precisión entre el conjunto de datos antes y después de los cambios paramétricos en el algoritmo de HOG.

```
> wilcox.test(before_accuracy, after_accuracy, paired = TRUE)

Wilcoxon signed rank test with continuity correction

data: before_accuracy and after_accuracy
V = 68, p-value = 0.002185
alternative hypothesis: true location shift is not equal to 0
```

**FIGURA 13.** Prueba de Wilcoxon para determinar diferencia significativa en la exactitud entre el conjunto de datos antes y después de los cambios paramétricos en el algoritmo de HOG.

```
> # Test de Wilcoxon para exactitud de CNN antes y después del ajuste
> wilcox.test(before_accuracy_cnn, after_accuracy_cnn, paired = TRUE)

Wilcoxon signed rank test with continuity correction

data: before_accuracy_cnn and after_accuracy_cnn
V = 60.5, p-value = 0.006278
alternative hypothesis: true location shift is not equal to 0
```

**FIGURA 14.** Prueba de Wilcoxon para determinar diferencia significativa en la exactitud entre el conjunto de datos antes y después de los cambios paramétricos en el algoritmo de CNN.

```
> # Test de Wilcoxon para precisión de CNN antes y después del ajuste
> wilcox.test(before_precision_cnn, after_precision_cnn, paired = TRUE)

Wilcoxon signed rank test with continuity correction

data: before_precision_cnn and after_precision_cnn
V = 24.5, p-value = 0.008303
alternative hypothesis: true location shift is not equal to 0
```

**FIGURA 15.** Prueba de Wilcoxon para determinar diferencia significativa en la precisión entre el conjunto de datos antes y después de los cambios paramétricos en el algoritmo de CNN.

## VIII. DISCUSIÓN

Los resultados obtenidos reflejan un avance significativo en la exactitud y precisión de los modelos de reconocimiento facial evaluados, particularmente bajo las condiciones controladas propuestas. Este análisis permite extraer conclusiones relevantes sobre el impacto de las mejoras implementadas y el desempeño de los modelos CNN y HOG en el contexto de sistemas de reconocimiento facial para entornos escolares.

### Desempeño de los Modelos CNN y HOG

El incremento en las métricas de exactitud y precisión demuestra que la optimización del sistema, a través de estrategias como la doble validación, el ajuste de parámetros de detección, la limitación del entrenamiento a imágenes específicas y el entorno de pruebas controlado, es altamente efectiva. En particular, la reducción de la tolerancia al 44% permitió una detección más precisa de rostros, eliminando falsos positivos (FP) sin comprometer los verdaderos positivos (VP).

Los resultados posicionan a ambos modelos con métricas casi idénticas: CNN con un 96% de exactitud y precisión, y HOG con 95% de exactitud y 94% de precisión. Esto representa una convergencia significativa entre los modelos, especialmente para HOG, que en estudios previos había mostrado limitaciones frente a CNN, con exactitud y precisión menores del 90%.

### Ventajas del Modelo HOG

HOG, conocido por su menor requerimiento computacional, destaca como una opción viable para instituciones educativas con recursos limitados, ahora que sus métricas casi igualan las de CNN. Esta optimización subraya la importancia de las mejoras implementadas, que permitieron a HOG competir directamente con CNN, tradicionalmente más robusto pero también más demandante en términos de hardware.

### Implicaciones para el Contexto Escolar

El entorno controlado, donde los estudiantes deben posicionarse frente a la cámara en condiciones de luz favorables, redujo significativamente las variables externas que anteriormente afectaban la detección. Esto refuerza la aplicabilidad del sistema en instituciones educativas, donde un entorno estandarizado es factible y deseable. Sin embargo, estas mismas condiciones podrían limitar el desempeño del sistema en entornos más dinámicos o menos controlados, lo que sugiere posibles líneas de investigación futura.

### Limitaciones y Consideraciones

Aunque los resultados son prometedores, es importante considerar que el sistema fue evaluado en condiciones ideales. La implementación en un entorno real podría presentar nuevos desafíos, como variaciones en el comportamiento de los usuarios o problemas técnicos en hardware menos potente que el utilizado en las pruebas (MacBook M3 Pro).

## IX. CONCLUSIÓN

Este estudio demuestra que las mejoras implementadas en el sistema de reconocimiento facial optimizaron significativamente el desempeño de los modelos CNN, alcanzando 96% de exactitud y 95% de precisión y HOG, alcanzando niveles de exactitud y precisión del 95% y 94%. Esto representa un avance notable en comparación con investigaciones previas, donde CNN y HOG presentaban un desempeño inferior bajo condiciones menos controladas.

Los resultados obtenidos por la prueba de Wilcoxon demuestran que las modificaciones al sistema (doble validación, segmentación de imágenes, y ajuste de tolerancia) tuvieron un impacto significativo en las métricas de rendimiento tanto para HOG como para CNN. En ambos casos, los valores de  $p$  fueron inferiores al nivel de significancia ( $\alpha = 0.05$ ), lo que permite rechazar la hipótesis nula y concluir que las mejoras implementadas fueron efectivas.

Además, el uso del test de Wilcoxon proporcionó una metodología robusta para evaluar la diferencia significativa entre las métricas antes y después de los ajustes, incluso ante la presencia de datos no normalmente distribuidos y la existencia de valores atípicos.

La implementación de estrategias como la doble validación, el ajuste de parámetros de detección y el entrenamiento limitado a imágenes específicas permitió maximizar el rendimiento de ambos modelos en un entorno diseñado para garantizar condiciones óptimas de iluminación y posicionamiento. Estos resultados no solo destacan la eficacia de las técnicas aplicadas, sino que también posicionan al modelo HOG como una alternativa viable y competitiva, especialmente en escenarios con restricciones de hardware, gracias a su menor demanda de recursos computacionales.

En el contexto escolar, el sistema presenta un potencial significativo para mejorar la gestión de la asistencia y otros procesos administrativos. Sin embargo, la dependencia de condiciones controladas para el reconocimiento sugiere la necesidad de explorar adaptaciones que permitan su aplicación en entornos menos predecibles. Adicionalmente, el uso de bases de datos limitadas resalta la importancia de validar el sistema en contextos más diversos y con poblaciones de usuarios más amplias.

El estudio no solo evidencia la efectividad de las estrategias aplicadas para optimizar el sistema, sino que también abre nuevas oportunidades para la investigación futura, enfocándose en la adaptabilidad y escalabilidad del sistema para satisfacer necesidades en entornos educativos y otros escenarios del mundo real.

## REFERENCIAS BIBLIOGRÁFICAS

- [1] FaceTime — Deep learning based face recognition attendance system. (n.d.). IEEE Conference Publication | IEEE Xplore. [https://ieeexplore.ieee.org/abstract/document/8080587?casa\\_token=4OBzR9HnbboAAAAA:j4utimT76x5s6ywfVozkm7HY9ji2ktgdoOLDF-pFrwo2aJvc9vve6PF9IA\\_MwIVoL\\_ZWF-GBJfHMRVkJ8](https://ieeexplore.ieee.org/abstract/document/8080587?casa_token=4OBzR9HnbboAAAAA:j4utimT76x5s6ywfVozkm7HY9ji2ktgdoOLDF-pFrwo2aJvc9vve6PF9IA_MwIVoL_ZWF-GBJfHMRVkJ8)
- [2] Automated Attendance System using OpenCV. (n.d.). IEEE Conference Publication | IEEE Xplore. <https://ieeexplore.ieee.org/abstract/document/9197936>
- [3] Aggarwal, M., Singh, R., Duggal, Y., et al. (2024). Modernizing Attendance Tracking: An Automated Headcount System Integrated with Face Detection. SpringerLink. [https://doi.org/10.1007/978-981-97-4149-6\\_34](https://doi.org/10.1007/978-981-97-4149-6_34)
- [4] Shukla, A. K., Shukla, A., & Singh, R. (2023). Automatic attendance system based on CNN-LSTM and face recognition. International Journal of Information Technology, 16(3), 1293–1301. <https://doi.org/10.1007/s41870-023-01495-1>
- [5] Comparative Study of Face Detection Using Cascaded Haar, HOG and MTCNN Algorithms. (2024). IEEE Xplore. <https://ieeexplore.ieee.org/document/10428242>
- [6] Emine, C., & Ahmet, C. (2017). Comparison of HOG (Histogram of Oriented Gradients) and Haar Cascade Algorithms with a Convolutional Neural Network Based Face Detection Approach. International Journal of Advance Research, Ideas and Innovations in Technology. <https://www.ijarjit.com>
- [7] Student attendance system in crowded classrooms using a smartphone camera. (n.d.). IEEE Conference Publication | IEEE Xplore. <https://ieeexplore.ieee.org/document/8658393>
- [8] Mamatkulovich, B. B. (2022, May 24). AUTOMATIC STUDENT ATTENDANCE SYSTEM USING FACE RECOGNITION. <http://nextscientists.com/index.php/science-conf/article/view/80>
- [9] Ishaq, K., & Bibi, S. (2023, August 3). IoT Based Smart Attendance System Using RFID: A Systematic Literature Review. arXiv.org. <https://arxiv.org/abs/2308.02591>
- [10] Sanivarapu, P. V. (2020). Multi-Face recognition using CNN for attendance system. In Lecture notes in networks and systems (pp. 313–320). [https://doi.org/10.1007/978-981-15-7106-0\\_31](https://doi.org/10.1007/978-981-15-7106-0_31)
- [11] Hasan, H. M., Rahman, M. M., Khan, M. A., Meghla, T. I., Mamun, S. A., & Kaiser, M. S. (2022). Implementation of Real-Time Automated Attendance System using deep learning. In Lecture notes in networks and systems (pp. 121–132). [https://doi.org/10.1007/978-981-16-7597-3\\_10](https://doi.org/10.1007/978-981-16-7597-3_10)
- [12] Hasan, H. M., Rahman, M. M., Khan, M. A., Meghla, T. I., Mamun, S. A., & Kaiser, M. S. (2022). Implementation of Real-Time Automated Attendance System using deep learning. In Lecture notes in networks and systems (pp. 121–132). [https://doi.org/10.1007/978-981-16-7597-3\\_10](https://doi.org/10.1007/978-981-16-7597-3_10)

