



**Módulo de software que implementa procesamiento digital de imágenes y técnicas de inteligencia artificial, para el análisis de imágenes en condiciones de baja luminosidad, para reconocimiento de presencia de personas**

Julián Ocampo Vélez

Informe de trabajo de grado para optar al título de Ingeniero Electrónico

Asesor:

David Fernández Mc Cann, PhD.

Universidad de Antioquia

Facultad de Ingeniería, Departamento de Ingeniería Electrónica

Pregrado

Medellín, Colombia

2024

<b>Cita</b>	(Ocampo Vélez, J. 2024)
<b>Referencia</b>	Ocampo Vélez, J. (2024). <i>Módulo de software que implementa procesamiento digital de imágenes y técnicas de inteligencia artificial, para el análisis de imágenes en condiciones de baja luminosidad, para reconocimiento de presencia de personas.</i> Trabajo de grado. Universidad de Antioquia, Medellín
<b>Estilo APA 7 (2020)</b>	



Centro de documentación de ingeniería (CENDOI)

**Repositorio Institucional:** <http://bibliotecadigital.udea.edu.co>

Universidad de Antioquia - [www.udea.edu.co](http://www.udea.edu.co)

**Rector:** John Jairo Arboleda Céspedes.

**Decano/Director:** Julio César Saldarriaga Molina.

**Jefe departamento:** Eduard Emiro Rodríguez Ramírez

El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Antioquia ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por los derechos de autor y conexos.

---

### **Agradecimientos**

Quiero agradecer primero que todo a mi familia, que hizo en gran parte lo posible para yo llegar a este paso en mi vida, a mis amigos, que siempre me apoyaron en todo momento, que estuvieron en momentos en los no sabía cómo seguir, en los que quería dejar este bonito proceso, y al final llegar a este último paso de un gran camino por recorrer como profesional, lo que es una gran satisfacción.

Agradecer a mi asesor David Fernández, quien me guio en mi proceso formativo como profesor en diferentes cursos y en la realización de este trabajo, y a las entidades y personas externas que confiaron en mí, para la realización de los objetivos de este trabajo.

Y, por último, pero no menos importante, a mí, por creer en que podía llegar a esta instancia, a confiar en mis capacidades, a aprender cosas que no creía en su momento y las experiencias que me llevo de todo mi proceso formativo.

---

## Resumen

Se presenta el desarrollo de módulo de software que utiliza técnicas de procesamiento digital de imágenes y algoritmos de inteligencia artificial para detectar la presencia de personas en condiciones de baja luminosidad.

El problema central abordado es la dificultad de identificar personas en imágenes con baja luminosidad, mediante redes neuronales convolucionales. El objetivo principal es crear un software capaz de mejorar la visibilidad y precisión en la detección de personas bajo estas condiciones.

El desarrollo del módulo se estructura en varias fases: un banco de pruebas diseñado en Python, en donde se simulan diferentes condiciones de iluminación degradando las imágenes de entrada, y la aplicación de métodos de procesamiento como la Ecuilización del Histograma Global (HE), Ecuilización de Histograma Adaptativo Limitado por Contraste (CLAHE) y Retinex Multi Escala con Restauración de Color (MSRCR). La detección de personas se lleva a cabo utilizando la red neuronal convolucional YOLO (You Only Look Once).

Los resultados muestran que HE es un método eficiente por su liviano. CLAHE, presenta una mejora local significativa, pero mantiene gran parte de la degradación de entrada y MSRCR es el método más robusto, en donde mejora significativamente la calidad de las imágenes y la precisión en la detección de personas en diversas condiciones de baja luminosidad. El estudio concluye que el módulo desarrollado puede ser una herramienta efectiva en el campo de la visión por computadora, seleccionando el procesamiento que más se adecue a las condiciones del escenario de implementación, con aplicaciones potenciales en la vigilancia y seguridad.

---

## Tabla de contenido

Agradecimientos .....	3
Resumen.....	4
Tabla de contenido .....	5
Lista de figuras.....	6
Lista de tablas .....	7
1. Introducción.....	8
2. Objetivos.....	11
2.1. Objetivo general.....	11
2.2. Objetivos específicos .....	11
3. Marco teórico.....	12
3.1. Ecualización del histograma global .....	12
3.2. Ecualización de histograma adaptativo limitado por contraste.....	13
3.3. Retinex .....	14
3.3.1. <i>Retinex de escala única</i> .....	15
3.3.2. <i>Retinex multi escala</i> .....	17
3.3.3. <i>Retinex multi escala con restauración de color</i> .....	17
3.4. Redes neuronales convolucionales .....	19
4. Metodología.....	21
4.1. Software de banco de pruebas.....	22
4.2. Imágenes degradadas .....	24
4.3. Procesamiento de imágenes .....	25
4.4. Detección de personas.....	27
5. Resultados y análisis.....	28
6. Conclusiones.....	41
Referencias.....	44

---

**Lista de figuras**

Figura 1 Histogramas de imágenes procesadas con HE y CLAHE .....	14
Figura 2 Diagrama de flujo para el desarrollo .....	22
Figura 3 Software desarrollado para banco de pruebas .....	23
Figura 4 Proceso para degradación de imágenes .....	25
Figura 5 Imagen original de entrada tomada para pruebas .....	29
Figura 6 Imágenes degradadas obtenidas del banco de pruebas .....	29
Figura 7 Salidas de imágenes procesadas y detección de personas con porcentaje de degradado de 0 .....	31
Figura 8 Salidas de imágenes procesadas y detección de personas con porcentaje de degradado de 20 .....	31
Figura 9 Salidas de imágenes procesadas y detección de personas con porcentaje de degradado de 50 .....	32
Figura 10 Salidas de imágenes procesadas y detección de personas con porcentaje de degradado de 80.....	33
Figura 11 Comparativa de los diferentes procesamientos .....	36
Figura 12 Histogramas.....	38

---

**Lista de tablas**

Tabla 1 Constantes para la implementación de Multiscale Retinex..... 17

Tabla 2 Resultados para los diferentes procesamientos en diferentes condiciones de degradado 34

---

## 1. Introducción

En escenarios cotidianos con entornos de poca iluminación, como días con niebla, lluvia o durante la noche, las imágenes que se capturan por medio de dispositivos electrónicos suelen presentar una variedad de degradaciones, lo que dificulta tanto la resolución humana como la visión por computadora y no se logran identificar las características de la imagen original [1, 2]. Por lo que, en entornos donde la identificación de presencia de personas es crucial, bajo estas condiciones se convierte en un desafío, y no se logra garantizar la efectividad y confiabilidad de los resultados que se obtienen al momento de la identificación y monitoreo, lo que representa una limitación al intentar aplicar estas tecnologías en casos como la videovigilancia.

Las condiciones de iluminación deficientes en distintos entornos suelen tener un impacto negativo en la calidad de las imágenes. En particular, las imágenes capturadas en condiciones de poca luz a menudo sufren problemas de bajo contraste, distorsión de color, degradación, entre otros, lo que conlleva a la pérdida de información y efectos visuales no deseados [3]. Aunque el uso de dispositivos profesionales o técnicas fotográficas avanzadas pueden mitigar en gran medida estas degradaciones, la falta de suficiente luz inevitablemente resulta en ruido y pérdidas de detalles y características en los resultados [4].

Las imágenes captadas por cámaras utilizadas para fines de seguridad en diferentes áreas a menudo no cumplen con los estándares de calidad deseados, ya sea debido a los costos asociados con cámaras de mayor calidad que podrían reducir estas degradaciones, o simplemente porque los dispositivos electrónicos son sensibles a diversos tipos de luz y fenómenos naturales como la lluvia o la niebla [5].

Debido a estas degradaciones persistentes en las imágenes, causadas por iluminación no uniforme y bajos contrastes, la información se enmascara o se pierde y realizar la detección de



---

presencia de personas se dificulta aún más, ya sea por visual humano o tecnología de visión por computadora. Esta limitación restringe las aplicaciones prácticas en el mundo real [2].

El problema que aborda este estudio surge de la necesidad de mejorar la detección de personas en condiciones de baja luminosidad, que es un desafío común en áreas como la seguridad y la vigilancia. A pesar de los avances en tecnologías de procesamiento de imágenes, la baja luminosidad sigue siendo un obstáculo significativo que impide una detección precisa y confiable.

El alcance del estudio incluye el desarrollo de un módulo de software que utiliza técnicas de procesamiento digital de imágenes e inteligencia artificial para mejorar la visibilidad y la detección de personas. Sin embargo, el estudio también reconoce ciertas limitaciones, como la necesidad de hardware especializado y la variabilidad en las condiciones de iluminación que pueden afectar la precisión del módulo.

En la actualidad, los modelos de inteligencia artificial, especialmente en el campo de las redes neuronales convolucionales, han logrado grandes avances en la detección de diferentes objetos o clases, según el entrenamiento específico de cada modelo. Sin embargo, en algunos casos, estos modelos se entrenan y utilizan condiciones óptimas de luminosidad y otros factores favorables que aseguran predicciones precisas en las imágenes de entrada, por lo que se hace crucial implementar un preprocesamiento que mitigue la limitante de que las imágenes de entrada no sean similares con las que fue entrada la red neuronal y así obtener un mayor rango de aplicación de los modelos de detección, dejando de importar el escenario en el que se encuentren las imágenes de entrada.

La metodología del estudio abarca varias etapas clave: la selección y análisis de tecnologías de detección de personas, la implementación de algoritmos de procesamiento de imágenes, y la validación del módulo mediante pruebas exhaustivas en diversos escenarios de baja luminosidad.

---

Se emplearon técnicas como la Ecuilización del Histograma, Ecuilización Adaptativa Limitada por Contraste, Retinex Multi Escala con Restauración de Color y redes neuronales convolucionales, para la detección de personas en las imágenes preprocesadas.

El estudio contribuye significativamente al campo del procesamiento digital de imágenes y la inteligencia artificial, ofreciendo una solución viable para la detección de personas en condiciones de baja luminosidad. Las aplicaciones prácticas de este trabajo son amplias, incluyendo mejoras que podrían ser implementadas en la seguridad y vigilancia, así como en la interacción humano-robot, proporcionando una herramienta avanzada y efectiva para enfrentar los desafíos de la baja luminosidad en estos contextos.

---

## 2. Objetivos

### 2.1. Objetivo general

Desarrollar módulo de software para detectar la presencia de personas en entornos de baja luminosidad, por medio de procesamiento digital de imágenes y técnicas de inteligencia artificial.

### 2.2. Objetivos específicos

- Realizar un análisis detallado de las tecnologías disponibles para la detección de personas en condiciones de iluminación baja, evaluando su idoneidad para su implementación en el módulo.
- Investigar y seleccionar algoritmos de procesamiento digital de imágenes y técnicas de inteligencia artificial adecuadas para la detección de personas en condiciones de iluminación baja.
- Desarrollar el módulo de software utilizando técnicas de Procesamiento Digital de Imágenes (PDI), con el fin de lograr una detección de personas con un nivel de precisión igual o superior al 70%. Esto implica la implementación de algoritmos de Inteligencia Artificial (IA), tales como redes neuronales convolucionales (CNN), que permitan una identificación en condiciones desafiantes de iluminación.
- Realizar pruebas exhaustivas y validaciones cruzadas, para garantizar la robustez y precisión del módulo en diversos escenarios y situaciones.

### 3. Marco teórico

#### 3.1. Ecuación del histograma global

Histogram Equalization (HE), es una de las técnicas más comunes para la mejora de contraste de una imagen con baja luminosidad. Básicamente, ajusta el histograma de la imagen original de tal manera que la distribución de niveles de grises sea más uniforme, esto ocasiona que se aumente el rango de tonos de gris en la imagen, lo que hace que los detalles sean más visibles y así mejorar el contraste general [1], dando como resultado una imagen con mejor luminosidad, resaltando detalles que no se logran identificar en la imagen original.

La ecuación consiste en transformar una distribución de datos, representada por un histograma, en otra distribución más uniforme y amplia de valores. Esto se hace para que los valores de los datos estén distribuidos de manera equitativa a lo largo de todo el rango disponible [2], donde el rango corresponde a la intensidad que cada valor de píxel puede tomar.

Primero, la imagen de entrada se divide en varios niveles de píxeles según su tono de gris, luego se calcula el histograma para cada nivel de tono de gris en los píxeles de la imagen, a los que se les calcula La Función de Distribución Acumulativa (CDF por sus siglas en inglés) y posteriormente sumando estos histogramas, por último, el valor de gris resultante después de la transformación se calcula utilizando un mapa de transformación de grises [2]. Se describe de la siguiente manera:

$$c(k) = \sum_{k=1}^i P(k) \quad (1)$$

Este procedimiento resulta en una imagen con mejor luminosidad y contraste; sin embargo, puede introducir ruido o distorsión no deseada debido a que aplica el proceso a toda la imagen,

---

incluyendo áreas que no requieren mejoras [2]. Este método es aplicable tanto en imágenes en escala de grises como en la ecualización de cada canal en un espacio de color, como por ejemplo RGB, en donde se realiza el mismo procedimiento, pero a cada canal, y posteriormente concatenando estos, obteniendo nuevamente la imagen en RGB con cada canal ecualizado según el procedimiento descrito para una imagen en escala de grises.

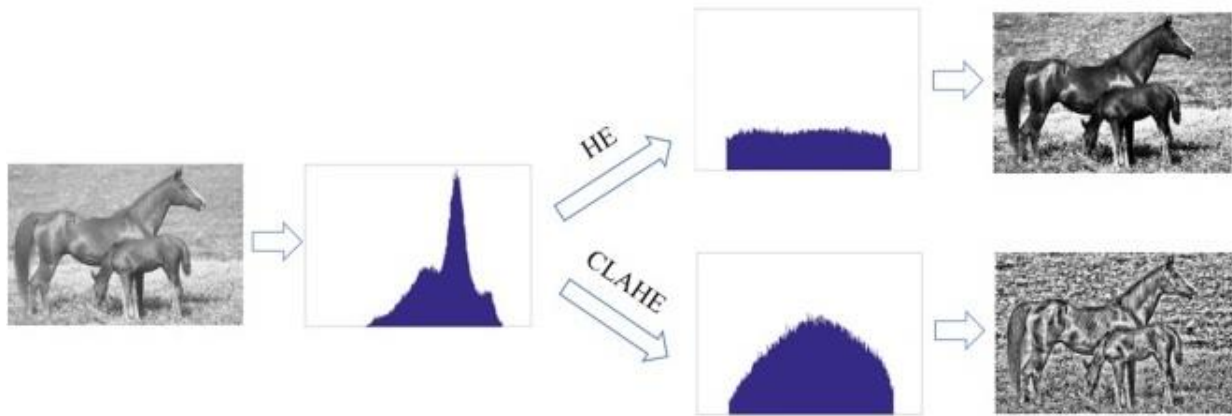
### **3.2. Ecualización de histograma adaptativo limitado por contraste**

Contrast-Limited Adaptive Histogram Equalization (CLAHE), resulta de una mejora para HE dado que, en ocasiones no cumple con los requisitos reales, ya que puede provocar que desaparezcan detalles o información en regiones que no necesiten mejoras u ocasionar ruido excesivo en la imagen, efecto no deseado que se quiere evitar cuando se busca mejorar la imagen final. Para abordar este problema se desarrolló CLAHE, que controla el aumento de contraste para reducir el efecto de amplificación del ruido [2].

CLAHE controla el aumento del ruido al recortar el histograma a un valor predeterminado, limitando y así la pendiente de la CDF. Este valor de recorte depende del tamaño de la región vecina y se distribuye uniformemente en todo el rango de grises para mantener la consistencia del histograma original [2]. Aplica localmente diversas transformaciones de escala de grises a cada región pequeña de la imagen, por lo que requiere la determinación del tamaño de la región, pero controlando la ecualización al limitar la ganancia de contraste local. Se han identificado dos inconvenientes de este método, la inevitable mejora del ruido en regiones suaves y la selección del límite de ganancia de contraste dependiente de la imagen [6].

**Figura 1**

*Histogramas de imágenes procesadas con HE y CLAHE*



*Nota.* Se observan los histogramas de salida al aplicar HE y CLAHE a una imagen en escala de grises, distribución del histograma y limitación de ganancia. Tomado de *A survey on image enhancement for Low-light images* 9(4), por Fu, X., Fan, G., Zhao, Y. & Wang, Z, (2023), *Heliyon*.

Como en el caso de HE, este procedimiento es aplicable a una imagen en RGB, o en definitiva a cualquier espacio de color, según sea la necesidad. Se aplica el procedimiento a cada canal individual y finalmente se concatenan nuevamente los tres canales, obteniendo como resultado una imagen ecualizada limitada por contraste en RGB.

**3.3. Retinex**

La palabra “Retinex”, es un acrónimo formado por “retina” y “cortex”, y se apoya en la premisa de que en la realidad los objetos no tienen color intrínseco, este surge de la interacción de la luz con ellos. Desarrollado por el Dr. Edwin Land, este enfoque emplea operaciones matemáticas que simulan el fenómeno de la constancia del color en el Sistema Visual Humano (HVS, por sus siglas en inglés). Cada tonalidad de color se genera a partir de una combinación de los colores primarios: rojo, verde y azul, cada uno con una longitud de onda específica. Estos tres colores

fundamentales son responsables de determinar la apariencia cromática de cada área individual [5]. Este modelo se fundamenta en la idea de que el color de un objeto se deriva de su capacidad para reflejar diferentes longitudes de onda de luz, como roja, verde y azul, en lugar de simplemente la cantidad total de luz reflejada. Logrando un equilibrio en el rango dinámico, mejora de bordes y manteniendo la invariancia de color, resultando en una mejora de la luminosidad [2, 5].

### 3.3.1. *Retinex de escala única*

Single-Scale Retinex (SSR), es una técnica perteneciente a la clase de funciones central/envolvente, donde la salida se obtiene al calcular la diferencia entre el valor de entrada (centro) y el promedio de su vecindad (entorno). En otras palabras, este método mejora el contraste y la apariencia visual de una imagen al comparar cada punto con un promedio de los valores circundantes [8].

Matemáticamente, la fórmula general para SSR se puede expresar, extendiendo la Ecuación 2 a todos los canales de color. La función toma en cuenta cada canal individualmente, aplicando la misma operación para procesar la imagen de manera integral. Esta técnica es particularmente útil en la corrección de imágenes degradadas, ya que resalta detalles y mejora la percepción del color [7].

$$R_i(x, y) = \log(I_i(x, y)) - \log(I_i(x, y) * F(x, y)) \quad (2)$$

donde  $I_i$  representa la imagen de entrada en el  $i$  –ésimo canal de color,  $R_i$  es la imagen de salida de Retinex en el  $i$  –ésimo canal, y  $F$  es la función envolvente normalizada. Esta operación se aplica de manera individual a cada canal de color, asegurando que cada uno sea procesado de forma independiente. La Ecuación 2 demuestra que el algoritmo mantiene la propiedad de "constancia

del color", en esencia, esto significa que el algoritmo es capaz de conservar la percepción del color de los objetos bajo diferentes condiciones de iluminación [8]. Tomando el supuesto que

$$I_i(x, y) = S_i(x, y) * r_i(x, y) \quad (3)$$

donde  $S_i$  es la iluminación y  $r_i$  es la reflectancia, se puede reescribir la Ecuación 2 como

$$R_i(x, y) \cong \log \frac{S_i(x, y) * r_i(x, y)}{\bar{S}_i * \bar{r}_i} \quad (4)$$

En la Ecuación 4, las barras indican el valor promedio ponderado, y se asume que  $S_i$  varía suavemente, siendo localmente casi constante. Por lo tanto, se puede aproximar  $S_i$  a  $\bar{S}_i$ . Así que

$$R_i \cong \log \frac{r_i}{\bar{r}_i} \quad (5)$$

Por lo tanto, se muestra que es independiente de la iluminación. Explicando así que el método atenúa los efectos de sombreado. Se propuso en su lugar una función gaussiana [8]

$$F(x, y) = C * \exp \left[ -\frac{x^2 + y^2}{2\sigma^2} \right] \quad (6)$$

en donde,  $\sigma$  representa la desviación estándar del filtro, la cual controla la cantidad de detalle espacial que se retiene y  $C$  siendo un factor de normalización tal que

$$\int F(x, y) dx dy = 1 \quad (7)$$

El valor de  $\sigma$  no puede ser modelado ni determinado teóricamente. Básicamente, existe un equilibrio entre la mejora de la dinámica local y la reproducción del color. Se encuentra que  $\sigma = 80$  [7], es una buena elección, encontrando este valor de una manera netamente experimental.



### 3.3.2. Retinex multi escala

Multiscale Retinex (MSR), surge debido al equilibrio necesario entre la compresión del rango dinámico y la reproducción del color. La elección adecuada de la escala de  $\sigma$  para el filtro envolvente  $F(x, y)$  es crucial en el método de SSR, por lo que MSR ofrece un equilibrio aceptable entre un buen rango dinámico local y una reproducción cromática adecuada. La salida de MSR se define como una suma ponderada de las salidas de varios SSR, por lo que propusieron que [7, 8]

$$R_{MSR_i} = \sum_{n=1}^N \omega_n R_{n_i} = \sum_{n=1}^N \omega_n [\log I_i(x, y) - \log(F_n(x, y) * I_i(x, y))] \quad (8)$$

donde  $N$  es el número de escalas,  $\omega_n$  es el peso de cada escala y  $F_n$  como la función definida en la Ecuación 6. Los valores de qué escalas elegir, cuáles y cuántas son los valores ideales que experimentalmente se definieron y se establecieron en las escalas de 15, 80 y 250 [8].

**Tabla 1**

Constantes para la implementación de Multiscale Retinex

Contante	$N$	$\sigma_1$	$\sigma_2$	$\sigma_3$	$\alpha$	$\beta$	$\omega_n$	$G$	$b$
Valor	3	15	80	250	125	46	1/3	192	-30

*Nota.* Esta tabla muestra los valores de constantes necesarios para la implementación del algoritmo de Multiscale Retinex. Tomada de *Multiscale Retinex, Image Processing On Line*, 4, pp. 71–88. por Petro, A B., Sbert, C. & Morel, J-M. (2014).

### 3.3.3. Retinex multi escala con restauración de color

Multiscale Retinex Color Restoration (MSRCR), propone un paso de restauración de color para complementar MSR [8], ya que, en ocasiones, se realizan ciertas suposiciones sobre la naturaleza general de los componentes de color en las imágenes. Una de estas suposiciones es la

del mundo gris, la cual establece que, dada una imagen con una cantidad suficiente de variaciones de color, el valor promedio de los componentes rojo, verde y azul debe corresponder a un valor de gris común [7]. En imágenes que no cumplen con esta suposición, es decir, aquellas donde un color específico puede dominar, el procedimiento Retinex anteriormente descrito produce imágenes grisáceas al disminuir la saturación del color.

Se sugiere modificar la salida de MSR multiplicándola por una función de restauración del color basada en la cromaticidad. El primer paso consiste en calcular las coordenadas de cromaticidad

$$I'_i(x, y) = \frac{I_i(x, y)}{\sum_{j=1}^S I_j(x, y)} \quad (9)$$

en donde para la  $i$  –ésima banda de color, donde  $S$  es el número de canales espectrales. (Generalmente,  $S = 3$  para el espacio de color RGB). El MSR de color restaurado viene dado por

$$R_{MSRCR_i}(x, y) = C_i(x, y) * R_{MSR_i}(x, y) \quad (10)$$

$$C_i(x, y) = f(I'_i(x, y)) \quad (11)$$

Donde para la  $i$  –ésima Función de Restauración de Color (CRF por sus siglas en inglés), encontraron que la mejor manera de definirla es [8]

$$C_i(x, y) = \beta \log[\alpha I'_i(x, y)] \quad (12)$$

donde  $\beta$  es una constante de ganancia y  $\alpha$  controla la fuerza de la no linealidad. Los autores [8] determinaron experimentalmente un único conjunto de valores para  $\beta$  y  $\alpha$  que parecen funcionar para todos los espectros (ver Tabla 1).

---

### 3.4. Redes neuronales convolucionales

El auge de la Inteligencia Artificial (IA) ha permeado todos los ámbitos de la vida, y la investigación en IA y Deep Learning (DL) ha logrado avances significativos. En la estructura de DL, el problema de la detección de objetos se aborda como la tarea de etiquetar distintos objetos en un cuadro de imagen con sus clases correctas y predecir sus cuadros delimitadores con alta probabilidad [9]. Los diferentes modelos de Redes Neuronales Convolucionales (CNN por sus siglas en inglés), facilitan la detección de diferentes objetos, animales, personas, entre otros, y en el que el reconocimiento de personas se considera un subconjunto específico de la tarea de detección en imágenes genéricas, con diversas aplicaciones potenciales, como la interacción humano-robot, la seguridad y la vigilancia [9].

Una de las CNN, que es aplicable a diferentes tareas, pero que funciona muy bien en el ámbito de detección de personas es YOLO (You Only Look Once) en sus diferentes versiones, donde se recomienda el uso de sus versiones Tiny, ya que es fácil de implementar en dispositivos de borde y con grandes capacidades de reconocimiento [3]. YOLO desarrollado por Joseph Redmon, Santosh Divvala, Ross Girshick y Ali Farhadi en 2016, emplea una arquitectura de CNN para identificar y localizar objetos de manera eficiente y precisa. La característica distintiva es que realiza la detección de animales, vehículos, alimentos o personas (como es el caso de interés), en una sola pasada a través de la red, a diferencia de otros enfoques que requieren múltiples pasadas, haciendo que YOLO sea rápido y adecuado para aplicaciones en tiempo real [10].

En la búsqueda de una detección óptima de objetos en tiempo real, YOLO en sus diferentes versiones funciona correctamente por su arquitectura, pero la versión 9 (la última al momento de realizar este trabajo) se destaca por su enfoque innovador para superar los desafíos de pérdida de información inherentes a las redes neuronales profundas. Este modelo integra la Información de

---

Gradiente Programable (PGI) y la arquitectura versátil de la Red de Agregación de Capas Eficiente Generalizada (GELAN). Estas integraciones no solo mejoran significativamente la capacidad de aprendizaje del modelo, sino que también aseguran la retención de información crucial a lo largo de todo el proceso de detección. Como resultado, YOLOv9 logra una precisión y un rendimiento excepcionales, estableciendo nuevos estándares en el campo de la detección de objetos en tiempo real. Esta versión presenta innovaciones a la hora de abordar los retos que se plantean a la hora de pérdida de la información en las redes neuronales convoluciones profundas, en la que garantiza una alta eficiencia y precisión, pese a ser un modelo simplificado, conserva y utiliza eficazmente la información esencial necesaria para la detección precisa de objetos [11].

#### 4. Metodología

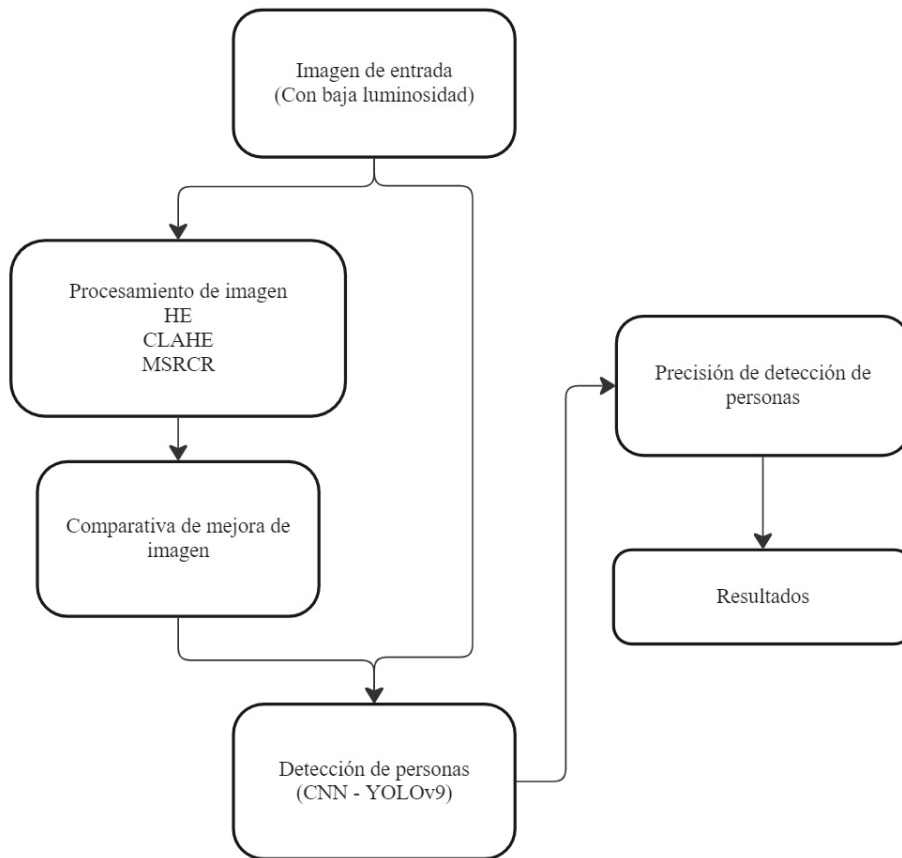
En esta sección se describe el procedimiento implementado para alcanzar los objetivos planteados en este trabajo. La Figura 2 muestra el flujo de desarrollo llevado a cabo, comenzando con el estudio de técnicas de preprocesamiento de imágenes como la Ecuación de Histograma, la Ecuación de Histograma Adaptativo limitado por contraste y Retinex Multi escala con Restauración de Contraste para mejorar el contraste y la luminosidad de la imagen degradada de entrada.

Para esto, se diseñó un software de banco de pruebas en donde cada imagen de entrada se le aplicó un procesamiento para degradarla y modificar su luminosidad, seguido de alguno de los tres preprocesamientos mencionados. Los resultados fueron evaluados tanto para mejorar la percepción visual humana como para optimizar la detección por computadora. Posteriormente, las imágenes resultantes fueron procesadas por Yolov9 para detectar la presencia de personas y medir la precisión en la detección, evaluando tanto la exactitud de detección como la cantidad de personas identificadas en cada imagen. Por último, se buscó encontrar un balance entre el tiempo necesario para el preprocesamiento y el tiempo requerido para la detección de personas en las imágenes, implementando finalmente un sistema de detección de personas efectivo en condiciones de baja luminosidad que puede ser implementado en videovigilancia.

Al final se logró una mejora de luminosidad y contraste de la imagen degradada, en la que la red neuronal convolucional fue capaz de identificar personas según sus pesos de entrenamiento. Lo que se buscó fue buscar la semejanza de imagen con la que fue entrada la red neuronal para lograr la detección de personas con un porcentaje de precisión aceptable.

**Figura 2**

*Diagrama de flujo para el desarrollo*



*Nota.* Las imágenes de entrada son obtenidas de un banco de pruebas y degradadas con la ayuda del software de banco de pruebas en diferentes escenarios de iluminación.

#### 4.1. Software de banco de pruebas

Se desarrolló un software en Python [12] utilizando diversas funcionalidades que ofrece este lenguaje. Se emplearon librerías como OpenCV [13] para el manejo de filtros y procesamiento de imágenes, y Flet [14] para el diseño de la interfaz gráfica.

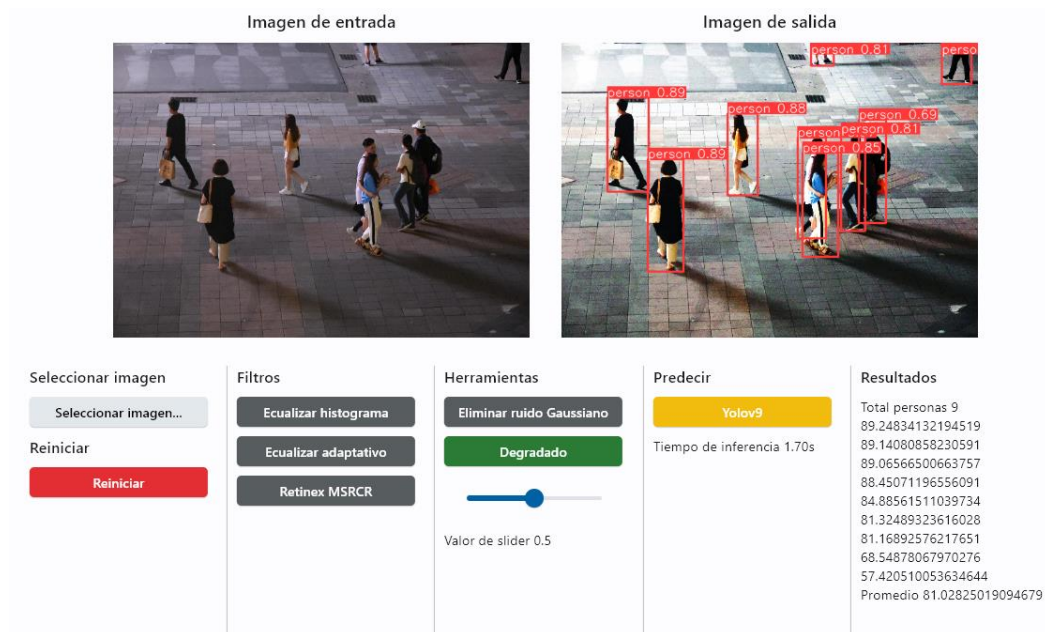
Consta de tres partes principales. La primera parte permite seleccionar uno de los procesamientos a evaluar, HE, CLAHE y MSRCR. La segunda parte denominada “herramientas”, ofrece la opción

de eliminar ruido gaussiano que se adiciona a la imagen luego del preprocesamiento, para evaluar posibles mejoras, y también se permite añadir un degradado uniforme a la imagen por medio de una barra deslizante. Por último, mediante YOLOv9 se predice la presencia de personas sobre la imagen que fue degrada y posteriormente procesada, encontrando el tiempo de inferencia y observando en la sección de resultados la cantidad de personas detectadas y su porcentaje de exactitud.

Para evaluar los diferentes filtros de mejora, se consideró calcular el promedio de los porcentajes de acierto, la cantidad de personas detectadas correctamente y el tiempo computacional requerido para aplicar las operaciones que conlleva cada procesamiento sobre la imagen de entrada.

**Figura 3**

*Software desarrollado para banco de pruebas*



---

*Nota.* Software desarrollado en funcionamiento, con una imagen de entrada, a la que se le aplica un degradado, uno de los procesamientos y finalmente la predicción para detectar la presencia de personas en la imagen.

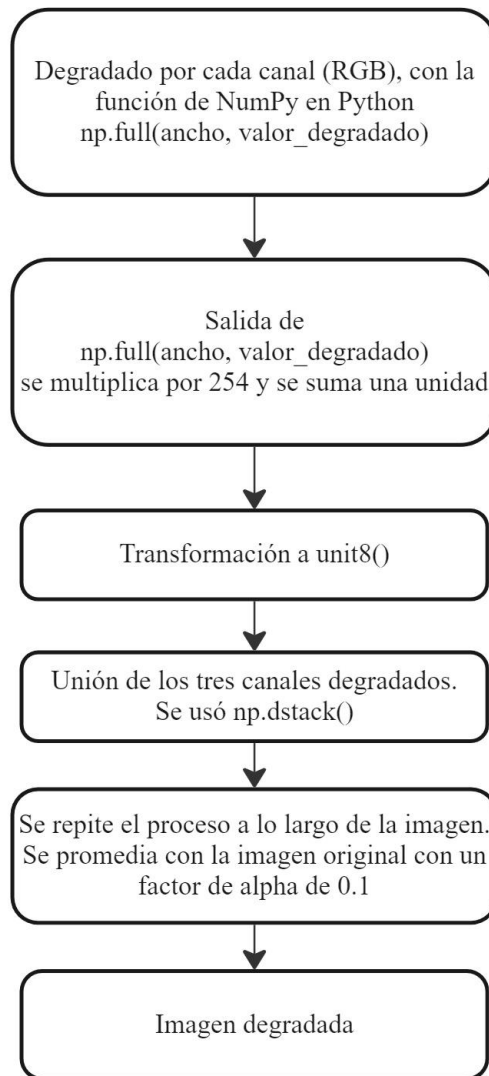
#### **4.2. Imágenes degradadas**

Para realizar este proceso, se añadió al software de banco de pruebas un control deslizante que permitió ajustar la luminosidad de las imágenes de entrada, haciéndolas más oscuras o claras. Este control facilitó la simulación de diferentes condiciones de iluminación a las que podrían estar expuestas las imágenes en situaciones reales. Al ajustar la luminosidad, se pueden evaluar cómo varían la calidad y la visibilidad de la imagen bajo distintos niveles de luz, desde entornos muy oscuros hasta aquellos extremadamente brillantes. Pese a estas degradaciones, se aseguró de que las imágenes mantuvieran su integridad y detalles en una variedad de escenarios. El proceso completo de ajuste y sus efectos en las imágenes se ilustran en la Figura 4 donde se presenta más detalladamente la estrategia implementada.



**Figura 4**

*Proceso para degradación de imágenes*



*Nota.* La figura describe el proceso para degradar las imágenes, para evaluar los resultados de los procesamientos y cómo es la mejora de estos ante condiciones adversas. Se implementó en el software de pruebas mediante la librería de NumPy.

### 4.3. Procesamiento de imágenes

A las imágenes degradadas, a las que se les modificó su luminosidad, desde muy oscuro hasta muy brillante, se les aplicó uno de los tres tipos de procesamientos: HE, CLAHE o MSRRCR,

---

donde cada uno de estos métodos tiene como objetivo mejorar la calidad visual de las imágenes a través de diferentes técnicas de ajuste y corrección.

El proceso de evaluación de la mejora obtenida mediante cada método se centró en dos aspectos principales: la luminosidad y el contraste. La luminosidad se refiere a la intensidad de luz en la imagen, mientras que el contraste se refiere a la diferencia de brillo entre las distintas áreas de la imagen. Una adecuada mejora en estos aspectos contribuyó significativamente a la claridad y calidad visual de la imagen final. Se observó que cada método de procesamiento presentó un rendimiento variable, con algunos comportamientos superiores a otros dependiendo de las condiciones específicas de la imagen de entrada.

Además, se examinó la eficacia de cada método de procesamiento en cuatro escenarios posibles de luminosidad, que incluyeron condiciones de baja luminosidad, alta luminosidad, luminosidad media, y luminosidad variable. Esta evaluación permitió determinar en qué condiciones específicas cada método proporciona una mayor mejora.

Otra parte fundamental del análisis se centró en observar cómo los valores de los píxeles se redistribuyeron a lo largo del histograma después de aplicar cada método de procesamiento, en donde el histograma de cada imagen es una representación gráfica de la distribución de los valores de brillo de los píxeles, y su análisis ofreció una visión detallada de cómo el procesamiento afecta la imagen en términos de contraste y detalle.

De esta manera, se buscó no solo mejorar la apariencia visual de las imágenes, sino también entender los mecanismos detrás de estas mejoras y las circunstancias bajo las cuales cada método es más efectivo. Este análisis integral proporciona una base sólida para elegir el método de procesamiento más adecuado para diferentes tipos de imágenes y condiciones de iluminación.

---

#### **4.4. Detección de personas**

Este proceso se llevó a cabo utilizando la red neuronal convolucional YOLO en su versión 9, se tomó la decisión de esta red ya que es de las últimas que se encontraban disponibles al momento de la elaboración de este trabajo y ofrece mejores resultados en comparación con sus versiones pasadas. Para implementar esta red, se utilizó un modelo pre entrenado con los pesos del conjunto de datos COCO (Common Objects in Context), que es ampliamente reconocido por su diversidad y volumen de imágenes etiquetadas de manera precisa.

En el estudio, cada imagen procesada fue ingresada a la CNN para identificar la cantidad de personas presentes. El rendimiento del modelo se evaluó al calcular el promedio de aciertos en la detección de personas en cada imagen. Este promedio de aciertos se refiere a la proporción de veces que el modelo identificó correctamente la presencia de personas en las imágenes comparado con la cantidad de personas reales presentes.

Este enfoque permitió no solo cuantificar la precisión del modelo de Yolov9 en la detección de personas, sino también identificar las fortalezas y debilidades del modelo bajo diferentes condiciones. Los resultados obtenidos proporcionaron una visión comprensiva de la capacidad del modelo para manejar imágenes mejoradas, y sirvieron como base para futuras mejoras y ajustes en el proceso de entrenamiento y optimización del modelo.

---

## 5. Resultados y análisis

Se presentan los resultados comparando los tres procesamientos de imagen presentados en este trabajo, aplicados sobre un dataset de imágenes que contaban en su estado original con una iluminación baja, y adicional un degradado, en donde se variaron los niveles de luminosidad desde muy oscuro hasta muy brillante, por medio del software de banco de pruebas desarrollado. El objetivo fue evaluar la efectividad de los tres métodos de procesamiento en mejorar la calidad de estas imágenes de entrada.

Para llevar a cabo este proceso, se utilizó un conjunto de datos de prueba que incluyó imágenes en condiciones de baja luminosidad como se muestra en la Figura 5 y a la que adicionalmente se degradó, evaluando cuatro posibles escenarios como se muestra en la Figura 7. Cada una de estas imágenes se sometió a uno de los tres métodos de procesamiento propuestos.

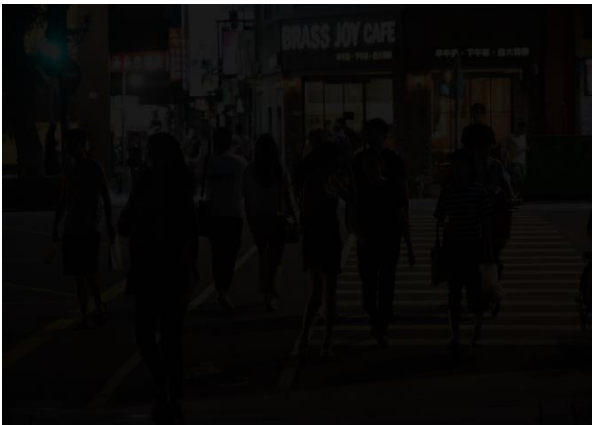
El análisis comparativo se centró en medir cómo cada método maneja los distintos niveles de luminosidad y el impacto resultante en la calidad de las imágenes, evaluando visualmente la mejora a nivel de luminosidad y contraste, y a nivel computacional al momento de realizar la detección de personas presentes en la imagen. Se tomó en cuenta de la totalidad de las personas presentes en la imagen, cuántas era capaz de detectar luego del procesamiento y con qué exactitud.

**Figura 5**

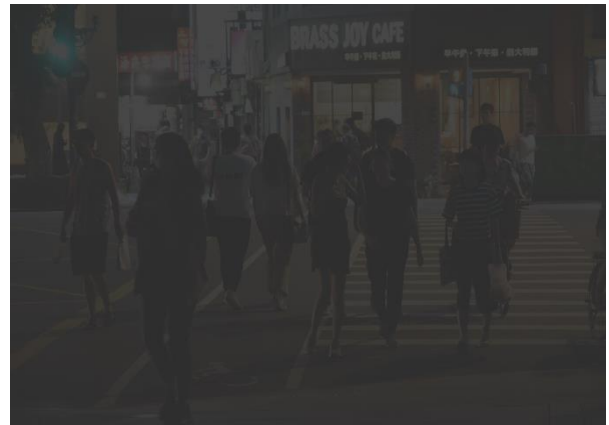
*Imagen original de entrada tomada para pruebas*

**Figura 6**

*Imágenes degradadas obtenidas del banco de pruebas*



a)



b)



c)



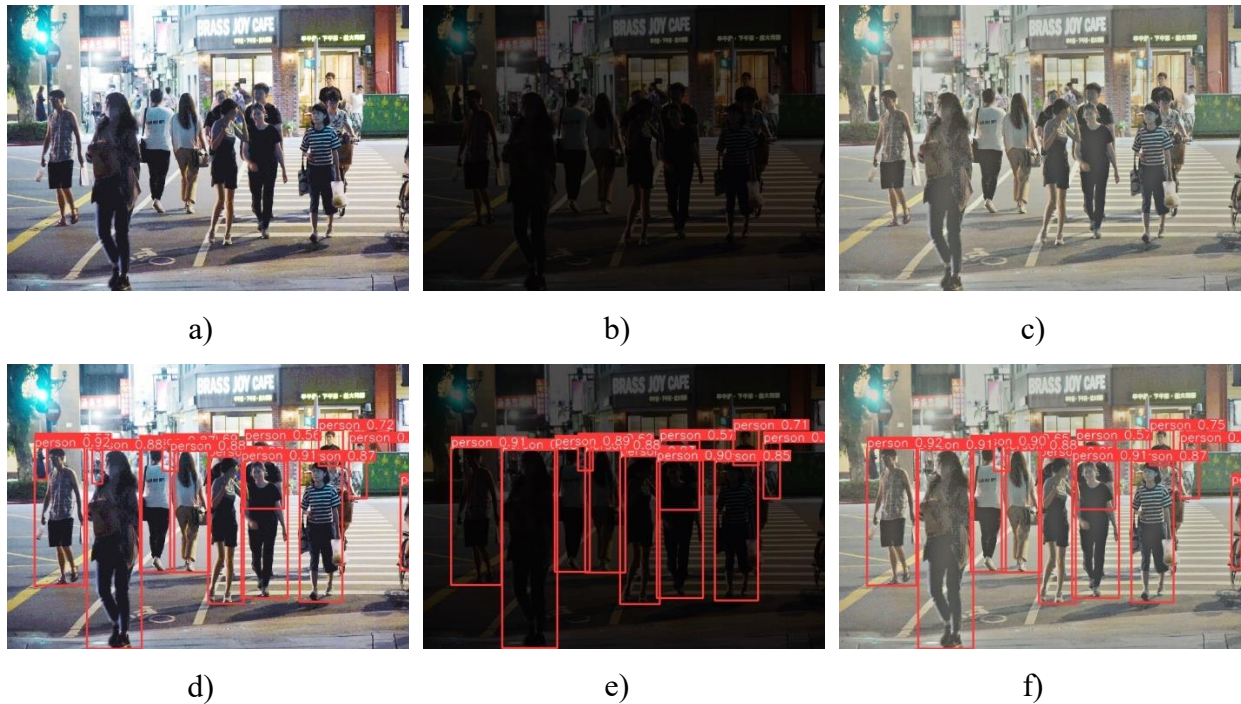
d)

*Nota.* Las diferentes imágenes corresponden a un porcentaje de degradado diferente, según la barra deslizante de degradado del banco de pruebas. a) Porcentaje de degradado de 0. b) Porcentaje de degradado de 20. c) Porcentaje de degradado de 50. d) Porcentaje de degradado de 80.

Se muestra el rendimiento de cada método de procesamiento ante los cuatro posibles escenarios propuestos. Se observa cómo cada método afecta las características de las imágenes de salida, proporcionando información de cómo es el comportamiento de cada uno de estos a nivel de mejora de iluminación y contraste. También se muestra cómo se evaluó la detección de personas en cada imagen de salida, mostrando cómo el mejorar las condiciones de luminosidad, con respecto a la imagen de entrada hacen que se obtengan resultados óptimos y una mejor precisión y visibilidad de los detalles de las imágenes.

**Figura 7**

*Salidas de imágenes procesadas y detección de personas con porcentaje de degrado de 0*



*Nota.* a) procesamiento con HE. b) Procesamiento con CLAHE. c) Procesamiento con MSRCR.  
 d) Detección de personas en imagen procesada con HE- e) Detección de personas en imagen procesada con CLAHE. f) Detección de personas en imagen procesada con MSRCR

**Figura 8**

*Salidas de imágenes procesadas y detección de personas con porcentaje de degrado de 20*







*Nota.* a) procesamiento con HE. b) Procesamiento con CLAHE. c) Procesamiento con MSRRCR. d) Detección de personas en imagen procesada con HE- e) Detección de personas en imagen procesada con CLAHE. f) Detección de personas en imagen procesada con MSRRCR

**Figura 9** Salidas de imágenes procesadas y detección de personas con porcentaje de degrado de 50



*Nota.* a) procesamiento con HE. b) Procesamiento con CLAHE. c) Procesamiento con MSRRCR. d) Detección de personas en imagen procesada con HE- e) Detección de personas en imagen procesada con CLAHE. f) Detección de personas en imagen procesada con MSRRCR



**Figura 10**

*Salidas de imágenes procesadas y detección de personas con porcentaje de degrado de 80*



*Nota.* a) procesamiento con HE. b) Procesamiento con CLAHE. c) Procesamiento con MSRCR. d) Detección de personas en imagen procesada con HE- e) Detección de personas en imagen procesada con CLAHE. f) Detección de personas en imagen procesada con MSRCR

Cada método fue puesto a prueba en los cuatro escenarios posibles, encontrando unas mejoras significativas a nivel visual humano, pero no tanto a nivel por visión por computadora en algunos casos, pero en otros fue el caso contrario, que a pesar, que para la visión humana no se encuentra una mejora, para el modelo de predicción sí, y predice mejor en comparativa a la imagen de entrada degrada o las otras imágenes de salida procesadas con los otros modelos de mejora.

Los resultados de las imágenes se recopilaban en la Tabla 2, donde se evaluaron varios aspectos clave del procesamiento: el tiempo necesario para completar el procesamiento de las imágenes, el tiempo de inferencia durante la detección de personas, el número de personas

identificadas y el promedio de precisión, si es este último aspecto utilizado como indicador del rendimiento general de cada método de procesamiento para mejorar la imagen.

Los procesamientos y diferentes resultados se tomaron en las mismas condiciones a nivel de hardware, tratando de tener siempre el mismo escenario de recursos computacionales, para tener unos mejores resultados comparativos entre cada procesamiento realizado, también cada imagen que se seleccionó del dataset, y luego de procesada antes de ser ingresada al modelo de predicción contaba con las mismas medidas de 480x640.

**Tabla 2**

Resultados para los diferentes procesamientos en diferentes condiciones de degradado

Porcentaje de degradado (%)	Imagen	Tiempo de procesamiento (ms)	Tiempo de inferencia (ms)	Número total de personas en la imagen	Porcentaje promedio de acierto (%)
—	Original	—	—	23	—
0	Degradada	—	1.63	7	70.83
	HE	3.03	1.52	14	74.75
	CLAHE	6.99	1.53	11	80.01
	MSRCR	4144.36	1.54	12	78.91
20	Degradada	—	1.66	10	79.04
	HE	3.02	1.71	14	74.83
	CLAHE	8.08	1.73	14	74.69
	MSRCR	4225.9	1.78	14	76.01
50	Degradada	—	1.67	9	82.43
	HE	1.99	1.67	13	76.84
	CLAHE	6.01	1.78	13	76.95
	MSRCR	3937.6	1.76	14	76.47

	Degradada	—	1.98	10	79.73
80	HE	2.01	1.78	14	75.25
	CLAHE	8.96	1.59	14	74.90
	MSRCR	4907.5	1.86	13	77.71

Se observó que en la imagen de prueba no hubo falsos positivos ni detecciones incorrectas que no fueran personas, lo cual es un aspecto importante que tiene el modelo de predicción utilizado, pero se tuvo en consideración que se puede dar el caso en los que se puedan tener predicciones erróneas, por lo que se debe tener precaución en ciertas situaciones. Cabe señalar que el modelo se limitó exclusivamente en la detección de personas, ya que YOLO, en su estructura, permite detectar muchas más clases. Por lo tanto, todas las personas detectadas fueron correctamente identificadas por el modelo.

Se encuentra que HE para la visión humana puede ofrecer buenos resultados, en la que algunos casos de pruebas, se observó que se resaltaban tanto los detalles como la iluminación de manera excesiva, que no mejoraba la imagen, sino que la empeoraba a nivel computacional, pero es un buen método, por ser un procesamiento liviano a comparación de CLAHE y MSRCR, y obtener muy buenos resultados para el enfoque que se le está dando.

CLAHE, se esperaban mejores resultados a nivel de visión humana, pero los resultados al momento de la predicción fueron buenos, incluso el modelo logró predecir más personas que con el procesamiento anterior.















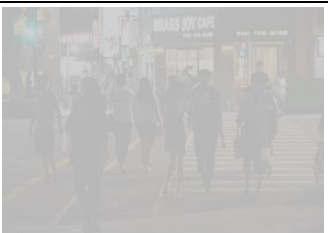

Para MSRCR, se encuentra que es un procesamiento robusto y pesado a nivel computacional por la base matemática que tiene, toma casi cuatro segundos en realizar la mejora

de imagen, pero los resultados son buenos tanto para la visual humana como para la visión por computadora.

A continuación, en la Figura 11, se comparan todas las salidas, luego de aplicar los respectivos procesamientos a la imagen de entrada degradada.

**Figura 11**

*Comparativa de los diferentes procesamientos*

Imagen degradada	Procesamiento de imagen		
	HE	CLAHE	MSRCR
			
			
			
			

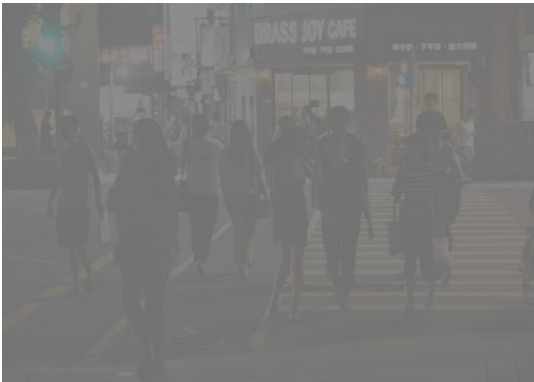
---

Para la visión humana, se observa que los procesamientos de HE y MSRCR mostraron mejores resultados al resaltar muy bien los detalles y mejorar significativamente el contraste. En el caso de HE, las imágenes muestran claridad en los detalles en comparación con la imagen original, aunque tienden a ser más brillantes debido a su enfoque en la ecualización global, que busca mejorar todo el rango de la imagen sin enfocarse tanto en los detalles, además, la saturación de los colores es notablemente mayor. Por otro lado, MSRCR mantiene una homogeneidad en los colores y los detalles se perciben de manera más coherente para la visión humana. En cuanto al método CLAHE, se observó que su capacidad de ecualización local limitada permitió resaltar detalles, aunque mantuvo cierto efecto de degradación lumínica inherente a la imagen original. CLAHE mostró esfuerzos notables en mejorar la claridad de detalles al operar a nivel local, lo cual es beneficioso para la percepción visual de ciertas regiones específicas de la imagen, sin embargo, la limitación en su capacidad para corregir completamente la iluminación degradada de entrada muestra que, aunque efectivo en áreas locales, aún puede beneficiarse de mejoras adicionales para lograr una corrección global más completa de la iluminación, como puede ser aplicar iterativamente el mismo procesamiento, pero no se determinó cuántas veces serían las suficientes, ya que depende mucho del estado de la imagen de entrada, y el objetivo es tener en cuenta, en lo posible, cualquier tipo de escenario con buenos resultados en la mayoría, no se evaluó cuántas veces serían suficientes aplicar el modelo sobre la imagen por su variabilidad con respecto al escenario en el que se encuentre la imagen de entrada degradada.

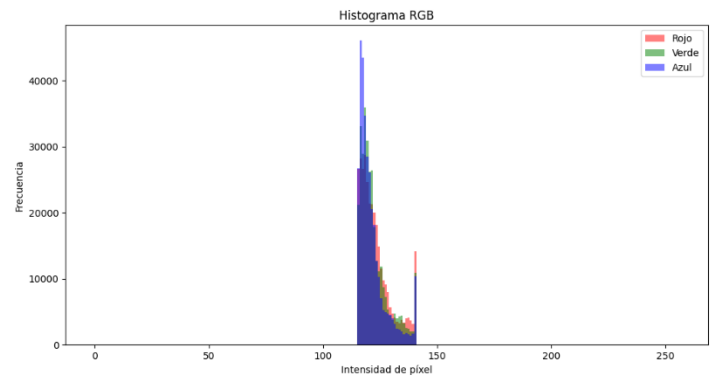
Se obtuvieron los histogramas, evaluando de esta manera como se amplía el rango de color en el espectro posible, en este caso de 0 a 255 que es la intensidad de cada píxel, siendo 0 oscuro y 255 claro. Se muestra el histograma de cada canal, rojo, verde y azul, en tanto en la imagen de entrada degrada como en las salidas que se obtuvieron al aplicar los tres métodos de procesamiento.

**Figura 12**

*Histogramas*



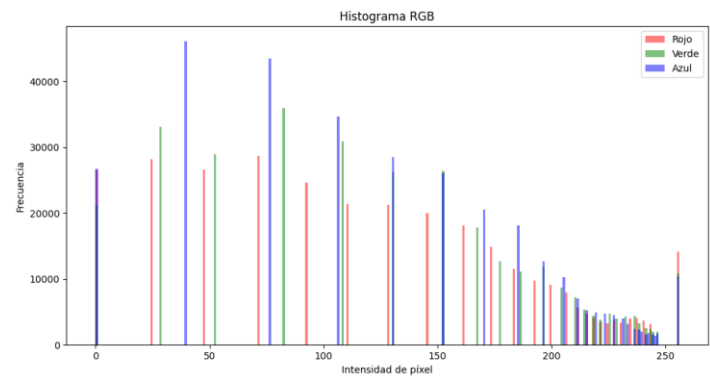
a)



b)



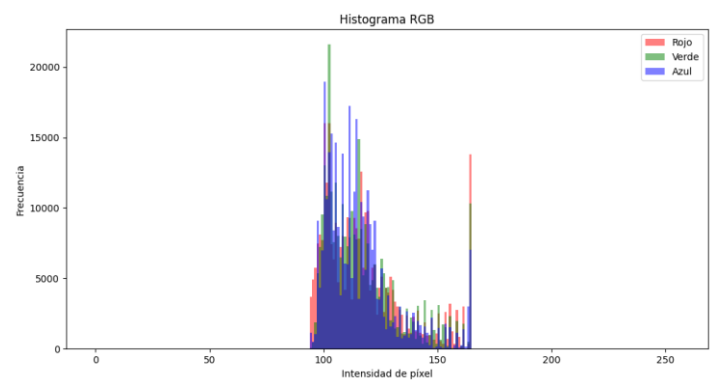
c)



d)



e)

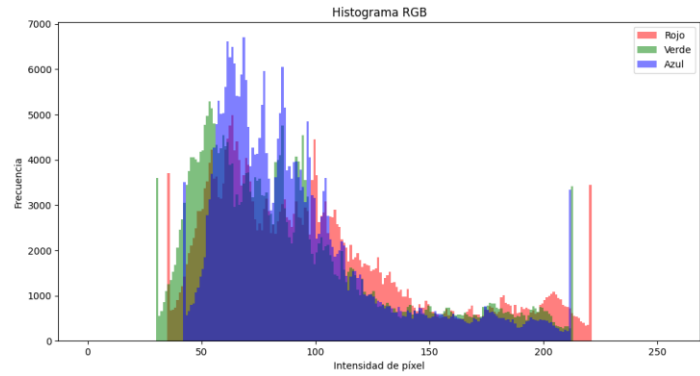


f)





g)



h)

*Nota.* Se muestran los diferentes histogramas de las imágenes de salida luego de aplicar el filtro de procesamiento a la imagen de entrada con un degradado de 50%. a) Imagen de entrada con degradado de 50%. b) Histograma de imagen de entrada. c) Imagen de salida procesada con HE. d) Histograma de imagen procesada con HE. e) Imagen de salida procesada con CLAHE. f) Histograma de imagen procesada con CLAHE. g) Imagen de salida procesada con MSRCR. h) Histograma de imagen de salida procesada con MSRCR.

Los histogramas muestran los resultados esperados para cada procesamiento, como se intuyó desde el marco teórico. Para el caso de HE, se encuentra que el rango de color se amplía en todo el rango posible sin importar en qué condiciones se encuentra la imagen de entrada, como se muestra en la Figura 11, las imágenes de salida para este caso son muy similares en los cuatro escenarios planteados.

Cuando se realiza el procesamiento con CLAHE, el espectro se estira un poco, pero limitado a su rango original, por lo que no se amplía muy dinámicamente, por lo que los resultados muestran que en la imagen de salida persistía, en cierta manera, la degradación inicial

Por último, en el caso de MSRCR, el rango se amplía mucho más, y más controlado, resaltando mejor los detalles y mejorando considerablemente la iluminación, por lo que es un

---

procesamiento muy bueno y robusto por la parte operacional, que conlleva llegar a este mucho más superior en comparativa a los otros dos procesamientos.

Los procesamientos revelaron resultados similares sin importar el escenario al que se sometían, teniendo en cuenta que para MSRCR y CLAHE, pueden mejorar debido a que tienen parámetros que se pueden modificar, se consideraron parámetros acordes a la literatura. Cada procesamiento presenta una mejora respecto a otra, ya sea a nivel visual humano o por visión por computadora, ya que, con respecto a la imagen de entrada degradada, la imagen de salida procesada claramente presenta una mejora en los detalles y el contraste. Según los resultados el mejor procesamiento, pese que para la visión humana no lo sea es CLAHE, por su tiempo de procesamiento y los resultados a la hora de la predicción comparada con los otros dos métodos. Si lo que se quisiera es una mejora completa sin importar el tiempo de procesamiento, claramente MSRCR tiene un mejor desempeño al momento de restaurar el color, resaltar los detalles y mejorar la luminosidad.



---

## 6. Conclusiones

- Se logró desarrollar un módulo de software que combina técnicas de procesamiento de imágenes, como la ecualización del histograma y el modelo Retinex, con redes neuronales convolucionales para la detección de personas. El sistema demuestra ser robusto en condiciones de baja luminosidad, mejorando significativamente la visibilidad y permitiendo una detección precisa de personas en imágenes degradadas.
- HE y CLAHE mostraron mejoras en la calidad de las imágenes, sin embargo, la técnica MSRCR resultó ser superior al proporcionar una mejora más controlada y destacando mejor los detalles en las imágenes, mejorando considerablemente la iluminación y contraste, tanto para la resolución humana como para la visión por computadora. MSRCR demostró ser altamente efectiva para el procesamiento de imágenes en condiciones de baja luminosidad. Esta técnica no solo amplió el rango de color de las imágenes, sino que también mejoró la visibilidad de detalles importantes, superando HE y CLAHE, pero se debe tener en cuenta el tiempo de procesamiento, ya que MSRCR tiene un tiempo considerable de procesamiento en comparación a los otros dos, por lo que se determina que CLAHE, pese a su limitante de conservar ciertos detalles de la imagen de entrada degradada, los resultados muestran que su tiempo de procesamiento y efectividad al momento de predecir la presencia de personas es superior a los otros dos métodos.
- La comparativa de los tres modelos queda a disposición, para una evaluación a nivel real, pero se muestra que se tienen precisiones aceptables en aproximadamente al 75% al momento de realizar la predicción con la red neuronal convolucional implementada, mostrando una mejoría con respecto a la imagen sin procesar.

- 
- La integración del modelo de detección de personas basado en redes neuronales convolucionales, específicamente utilizando el modelo YOLO, permitió una detección precisa y eficiente de personas en las imágenes mejoradas. La evaluación demostró que el modelo es capaz de detectar personas con alta precisión incluso en condiciones de baja luminosidad, debido al preprocesamiento realizado con los tres modelos planteados.

Departamento de Ingeniería Electrónica y de Telecomunicaciones

Módulo de software que implementa procesamiento digital de imágenes y técnicas de inteligencia artificial, para el análisis de imágenes en condiciones de baja luminosidad, para reconocimiento de presencia de personas



UNIVERSIDAD DE ANTIOQUIA

Facultad de Ingeniería

PRACTICANTE: Julián Ocampo Vélez

ASESORES: David Stephen Fernández Mc Cann

PROGRAMA: Ingeniería electrónica

Semestre de la práctica: 2024-1

Este trabajo se enfocó en el desarrollo de un software que simula diferentes escenarios de luminosidad y utilizando técnicas de procesamiento digital de imágenes realizar la mejora de luminosidad de imágenes y, a través de redes neuronales convolucionales, detectar la presencia de personas en las imágenes preprocesadas.



### Introducción

En condiciones de baja iluminación, como niebla, lluvia o noche, las imágenes capturadas por dispositivos electrónicos se degradan, dificultando la identificación de detalles ya sea por visión humana o por computadora. Esto es un reto en aplicaciones de seguridad y videovigilancia. El estudio aborda este problema desarrollando un módulo de software que utiliza técnicas de procesamiento digital de imágenes para mejorar la visibilidad e implementa redes neuronales convolucionales para realizar la detección de personas. A pesar de los avances tecnológicos, la baja luminosidad sigue siendo un obstáculo significativo, y este trabajo busca mitigar sus efectos mediante diversas técnicas, incluyendo ecualización de histograma y modelos Retinex, para el preprocesamiento de las imágenes y el modelo de detección de objetos, limitado a solo personas YOLOv9.



### Objetivos

- ✓ Analizar tecnologías para la detección de personas en condiciones de baja luminosidad, evaluando su idoneidad para su implementación en el módulo.
- ✓ Investigar y seleccionar algoritmos de procesamiento digital de imágenes y técnicas de inteligencia artificial adecuadas para la detección de personas en condiciones de iluminación baja.
- ✓ Desarrollar el módulo de software utilizando técnicas de Procesamiento Digital de Imágenes, logrando la detección de personas con precisión igual o superior al 70%. Implementando algoritmos de Inteligencia Artificial, como redes neuronales convolucionales, que permitan una identificación en condiciones desafiantes de iluminación.
- ✓ Realizar pruebas exhaustivas y validaciones cruzadas, para garantizar la robustez y precisión del módulo en diversos escenarios y situaciones.



### Metodología

Investigación de diversas técnicas para mejorar la luminosidad en imágenes. Debido a su nivel de complejidad y eficacia, se seleccionaron tres métodos principales: ecualización global de histograma, ecualización de histograma adaptativa limitada por contraste y Retinex multi escala con restauración de color. Elección del modelo YOLO. Implementación de software en Python y validación de eficacia y precisión.

Imagen degradada	Procesamiento de imagen		
	HE	CLAHE	MSRCR



### Resultados

Los resultados indican que cada uno de los métodos de preprocesamiento presenta sus propias ventajas, lo que permite seleccionar el más adecuado según el entorno de implementación. Todos los métodos de preprocesamiento mostraron una mejora significativa tanto en la calidad visual humana como en el análisis computacional, facilitando que la red neuronal convolucional detecte la presencia de personas con mayor precisión en comparación con las imágenes sin preprocesar. En los cuatro escenarios de evaluación, los tres métodos de preprocesamiento mostraron resultados similares, independientemente del escenario.



### Conclusiones

- ✓ Se desarrolló módulo de software que integra técnicas de procesamiento de imágenes y redes neuronales convolucionales para la detección de personas. Este sistema es eficaz en condiciones de baja luminosidad.
- ✓ HE y CLAHE mostraron mejoras significativas pero MSRCR, mostró una mejora mas detallada de color y contraste, mejorando tanto la visión humana como por computadora.
- ✓ La precisión fue aproximadamente del 75% en los tres métodos, pero fue superior con las imágenes preprocesadas con MSRCR.
- ✓ El modelo de YOLOv9, muestra una alta eficiencia para la detección de personas con tiempos de inferencia adecuados para la implementación.

DATOS DE CONTACTO DEL AUTOR:

3117500615

+57 3117500615

julian.ocampov@udea.edu.co

https://www.linkedin.com/in/julián-ocampo-vélez/

---

### Referencias

- [1] Fu, X., Fan, G., Zhao, Y. & Wang, Z. (2011). A new image enhancement algorithm for low illumination environment. *IEEE International Conference on Computer Science and Automation Engineering*. 625-627. DOI: 10.1109/CSAE.2011.5953297.
- [2] Guo, J., Ma, J., García-Fernández, A. F., Zhang Y., & Liang, H. (2023). A survey on image enhancement for Low-light images. *Heliyon*, 9(4). DOI: <https://doi.org/10.1016/j.heliyon.2023.e14558>.
- [3] Hussain, A., Khan, S. U., Khan, N., Rida, I., Alharbi, M., & Baik, S. W. (2023). Low-light aware framework for human activity recognition via optimized dual stream parallel network. *Alexandria Engineering Journal*, 74, 569-583. <https://doi.org/10.1016/j.aej.2023.05.050>.
- [4] Li, M., Liu, J., Yang, W., Sun, X. & Guo, Z. (2018). Structure-Revealing Low-Light Image Enhancement Via Robust Retinex Model. *IEEE Transactions on Image Processing*, 27(6). 2828-2841. 10.1109/TIP.2018.2810539.
- [5] Rosales, A J., Gallegos, A J., Dehesa, M., Souvrevile, S. & Hernández, I. (2015). Implementación del modelo Retinex aplicado al procesamiento de imágenes subacuáticas para mejorar su contenido cromático. *Research in Computing Science* 91(1):95-103. 10.13053/rcs-91-1-8.
- [6] Bassiou, N. & Kotropoulos, C. (2007). Color image histogram equalization by absolute discounting back-off. 107(1-2), 108-122. DOI: 10.1016/j.journal.2007.01.001.
- [7] Petro, A B., Sbert, C. & Morel, J-M. (2014). Multiscale Retinex, *Image Processing On Line*, 4, pp. 71–88. <https://doi.org/10.5201/ipol.2014.107>.

- 
- [8] Jobson, D J., Rahman, Z. & Woodell, G A. (1997). Properties and performance of a center/surround retinex, IEEE Transactions.
- [9] Sivachandiran, S., Mohan, K J. & Nazer, G M. (2022). Deep learning driven automated person detection and tracking model on surveillance videos. *Measurement: Sensors*, 24, 100422. <https://doi.org/10.1016/j.measen.2022.100422>.
- [10] Solawetz, J. (2024). What is YOLOv7? A Complete Guide. Recuperado de: <https://blog.roboflow.com/yolov7-breakdown/>.
- [11] Ultralytics. (s.f.). *YOLOv9 models*. Ultralytics. Recuperado en julio de 2024, de <https://docs.ultralytics.com/es/models/yolov9/>.
- [12] Python Software Foundation. (s.f.). Python programming language. Recuperado en junio de 2024, de <https://www.python.org/>.
- [13] OpenCV. (s.f.). Recuperado en junio de 2024, de <https://opencv.org/>.
- [14] Flet Documentation. (s.f.). Recuperado en junio de 2024, de <https://flet.dev/docs/>.