



Tecnológico Nacional de México

Centro Nacional de Investigación
y Desarrollo Tecnológico

Tesis de Maestría

Reconocimiento automático del rostro para
verificación de identidad para evaluación en línea

Presentada por
Ing. Wendy Valderrama Cardenas

Como requisito para la obtención del grado de
Maestra en Ciencias de la Computación

Directora de tesis
Dra. Andrea Magadán Salazar

Cuernavaca, Morelos, México. Junio de 2019.



SEP
SECRETARÍA DE
EDUCACIÓN PÚBLICA



TECNOLÓGICO NACIONAL DE MÉXICO

Centro Nacional de Investigación y Desarrollo Tecnológico

"2019, Año del Caudillo del Sur, Emiliano Zapata"

Cuernavaca, Morelos a 06 de junio del 2019
OFICIO No. DCC/028/2019

Asunto: **Aceptación de documento de tesis**

DR. GERARDO V. GUERRERO RAMÍREZ
SUBDIRECTOR ACADÉMICO
PRESENTE

Por este conducto, los integrantes de Comité Tutorial de la **Ing. Wendy Valderrama Cardenas**, con número de control M17CE103, de la Maestría en Ciencias de la Computación, le informamos que hemos revisado el trabajo de tesis profesional titulado "**Reconocimiento automático del rostro para verificación de identidad para evaluación en línea**" y hemos encontrado que se han realizado todas las correcciones y observaciones que se le indicaron, por lo que hemos acordado aceptar el documento de tesis y le solicitamos la autorización de impresión definitiva.

DIRECTORA DE TESIS

Dra. Andrea Magadán Salazar
Doctorado en Ciencias
Computacionales
10654097

REVISOR 1

M.C. Gerardo Reyes Salgado
Maestro en Ciencias de la
Computación
2493370

REVISOR 2

Dr. José Ruiz Ascencio
Doctor en Ciencias
5009035

C.p. M.T.I. M.E. Guadalupe Garrido Rivera - Jefa del Departamento de Servicios Escolares.
Estudiante
Expediente

NACS/lmz



SEP
SECRETARÍA DE
EDUCACIÓN PÚBLICA



TECNOLÓGICO NACIONAL DE MÉXICO

Centro Nacional de Investigación y Desarrollo Tecnológico

"2019, Año del Caudillo del Sur, Emiliano Zapata"

Cuernavaca, Mor., 6 de junio de 2019
OFICIO No. SAC/215/2019

Asunto: Autorización de impresión de tesis

ING. WENDY VALDERRAMA CARDENAS
CANDIDATA AL GRADO DE MAESTRA EN CIENCIAS
DE LA COMPUTACIÓN
P R E S E N T E

Por este conducto, tengo el agrado de comunicarle que el Comité Tutorial asignado a su trabajo de tesis titulado "Reconocimiento automático del rostro para verificación de identidad para evaluación en Línea", ha informado a esta Subdirección Académica, que están de acuerdo con el trabajo presentado. Por lo anterior, se le autoriza a que proceda con la impresión definitiva de su trabajo de tesis.

Esperando que el logro del mismo sea acorde con sus aspiraciones profesionales, reciba un cordial saludo.

ATENTAMENTE

Excelencia en Educación Tecnológica®

"Conocimiento y tecnología al servicio de México"

DR. GERARDO VICENTE GUERRERO RAMÍREZ
SUBDIRECTOR ACADÉMICO



SEP TecNM
CENTRO NACIONAL
DE INVESTIGACIÓN
Y DESARROLLO
TECNOLÓGICO
SUBDIRECCIÓN
ACADÉMICA

C.p. Mtra. Guadalupe Garrido Rivera.- Jefa del Departamento de Servicios Escolares.
Expediente

GVGR/mcr

cenidet[®]
Centro Nacional de Investigación
y Desarrollo Tecnológico

Interior Internado Palmira S/N, Col. Palmira, C. P. 62490, Cuernavaca, Morelos.
Tel. (01) 777 3 62 77 70, ext. 4106, e-mail: dir_cenidet@tecnm.mx

www.tecnm.mx | www.cenidet.edu.mx



Dedicatoria

A Dios, Victoria, José y Mirsha

Agradecimiento

A Dios por permitirme cumplir este sueño.

A mis padres Victoria y José, por su apoyo incondicional, sus palabras de aliento en los buenos y malos momentos que me ayudan a ser un mejor ser humano y seguir creciendo como persona.

A Mirsha por ayudarme siempre que lo necesito y alentarme a seguir adelante cumpliendo mis metas.

A mi asesora la Dra. Andrea Magadan Salazar por su paciencia, su tiempo, sus valiosos consejos y sobre todo el gran apoyo brindado.

A mi comité revisor el Dr. José Ruíz Ascencio y el Dr. Gerardo Reyes Salgado por sus comentarios y observaciones que me ayudaron a culminar este proyecto.

A mis compañeros por sus recomendaciones y apoyo.

A mis compañeros y amigos que me ayudaron a la creación de mi banco de imágenes, por brindarme ese espacio y tiempo.

A el Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET) por las facilidades y el apoyo proporcionado para la terminación de esta tesis.

A el Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo económico que me brindó durante mis estudios de maestría.

Finalmente agradezco a todos los que de forma directa o indirecta contribuyeron a la realización de este proyecto.

Resumen

En este proyecto de tesis se desarrolló un sistema de visión artificial que monitorea, en tiempo real, a los aspirantes que presentan el examen de admisión, mientras presentan dicho examen en línea; esto a través de capturas de imágenes de su cámara web. Es decir, durante este tiempo, el sistema realiza capturas del rostro de los aspirantes para verificar su identidad; se pretende comprobar que la persona que se registró en el sistema con su credencial del Instituto Nacional Electoral (INE) para presentar el examen, sea la misma que se encuentra frente a la computadora durante la ejecución del mismo.

Este sistema permite proveer una opción adicional para garantizar la seguridad en exámenes de admisión en línea, logrando que los aspirantes no realicen suplantación de identidad y reducir la carga de supervisión manual del examen.

Abstract

This thesis project presents an artificial vision system that monitors the applicants in real time while they are taking their online exam. This is carry out by capturing images from their webcam. That is, during this time the system captures the face of the aspirants to verify their identity; it is intended to verify that the person who was registered in the system with his/her credential of the National Electoral Institute (INE) at the moment of presenting the exam is the same person that it is in front of the computer.

This system allows to provide an additional option to guarantee security of online admission exams, making sure that applicants do not perform identity theft and reducing the manual supervision load of the exam.

Índice

Resumen.....	I
Abstract	II
Lista de Figuras.....	V
Lista de Tablas	VIII
Glosario	IX
CAPÍTULO 1 Introducción	1
1.1 Planteamiento del problema	2
1.2 Complejidad del problema	2
1.3 Objetivo general.....	2
1.3.1 Objetivos específicos.....	2
1.4 Alcances y limitaciones.....	3
1.5 Metodología de solución.....	3
1.6 Organización de la tesis.....	4
CAPÍTULO 2 Estado del arte	6
2.1 Reconocimiento facial	7
2.2 Reconocimiento facial con iluminación variable.....	12
2.3 Análisis del estado del arte	19
2.4 Estado de la práctica	19
CAPÍTULO 3 Marco teórico.....	22
3.1 Técnicas de pre-procesamiento	23
3.2 Algoritmo propuesto: foto receptores con retinex multi-escala (FRRM)	25
3.3 Descripción del rostro	27
3.4 Algoritmos de clasificación.....	28
3.5 Métricas.....	29
CAPÍTULO 4 Diseño del sistema	32
4.1 Funcionamiento del sistema	33
4.2 Descripción del rostro	35
4.3 Interface del sistema	37
4.4 Herramientas utilizadas (software y equipo).....	39
CAPÍTULO 5 Pruebas y Resultados	40
5.1 Banco de imágenes	41

5.2	Experimentación	43
5.3	Análisis de resultados.....	44
5.3.1	Localización	44
5.3.2	Pre-procesamiento.....	47
5.3.3	Verificación de identidad	55
	Tiempo	62
5.4	Comentarios finales.....	64
CAPÍTULO 6 Conclusiones		65
6.1	Conclusiones.....	66
6.2	Objetivos logrados.....	67
6.3	Aportaciones	68
6.4	Trabajo futuro	68
6.5	Productos académicos adicionales	69
Anexos		70
A.	Manual de instalación (OpenCV [41], Dlib [42] y Qt [45]).....	70
B.	Productos académicos adicionales	80
Referencias.....		84

Lista de Figuras

Figura 1.1 Esquema de metodología de solución.....	3
Figura 2.1 (a) Imágenes con cambios de expresión e iluminación. (b) imágenes con diferentes tipos de oclusiones (disfraz real y oclusión de bloques aleatorios). La tarea de reconocimiento en estas imágenes (b) es mucho más difícil que la tarea de reconocimiento que en (a) [4]	8
Figura 2.2 Categorías de algoritmos de reconocimiento de rostros tradicionales. [5]	9
Figura 2.3 Comparación visual de imágenes faciales de la base de datos Yale face database B, después de aplicar diferentes métodos [10].....	12
Figura 2.4 Diagrama del proceso de funcionamiento del programa Open Face [29].....	19
Figura 3.1 Valor de 3x3 píxeles de la porción del ojo izquierdo antes y después de aplicar el histograma [16].....	23
Figura 3.2 (a) Imagen en escala de grises, (b) ecualización del histograma tradicional, (c) ecualización adaptativa. [33]	24
Figura 3.3 a) Imagen de entrada con mala iluminación, b) imagen con buena iluminación. La imagen de salida muestra el resultado de aplicar la capa de foto receptores a la imagen de entrada.	26
Figura 3.4 Resultado visual de la aplicación de los algoritmos listados: a) Imagen de entrada, b) retina, c) retinex multi-escala, d) técnica propuesta FRRM.....	27
Figura 3.7 Objetivo del algoritmo de clasificación en el sistema.....	29
Figura 3.8 Estructura general de la matriz de confusión [39].....	30
Figura 4.1 Diagrama del proceso del sistema.....	33
Figura 4.2 Puntos preestablecidos del rostro [40].....	36
Figura 4.3 a) Puntos de interés del rostro, b) distancias faciales obtenidas a partir de los puntos de interés.	36
Figura 4.4 Interfaz de entrenamiento del sistema.....	38
Figura 4.5 Interfaz de monitoreo del sistema.....	39
Figura 5.1 Variaciones de imágenes por individuo del banco de imágenes FEI [46]	41
Figura 5.2 Muestra de imágenes contenidas en el banco de imágenes Faces94. [47] ...	42
Figura 5.3 Muestra de las imágenes contenidas en el banco de imágenes propio.....	43
Figura 5.4 (a) Imágenes con iluminación controlada, (b) imágenes con iluminación variable.	47
Figura 5.5 Imágenes resultantes de aplicar el filtro de ecualización del histograma a imágenes con iluminación controlada e iluminación variable.	48
Figura 5.6 Imágenes resultantes de aplicar el filtro de ecualización adaptativa del histograma a imágenes con iluminación controlada e iluminación variable.	48
Figura 5.7 Imágenes resultantes de aplicar el filtro de modelo de retina a imágenes con iluminación controlada e iluminación variable.....	49
Figura 5.8 Imágenes resultantes de aplicar el filtro de retinex multi-escala a imágenes con iluminación controlada e iluminación variable.....	49
Figura 5.9 Imágenes, con iluminación controlada e iluminación variable, resultantes de haberles aplicado la técnica propuesta FRRM.	50

Figura 5.10 Imágenes con retina + reinex aplicado a imágenes con iluminación controlada e iluminación variable.	50
Figura 5.11 Gráfica comparativa de algoritmos de pre-procesamiento respecto a la métrica SSIM.....	51
Figura 5.12 Imágenes con iluminación controlada e iluminación variable, resultantes de aplicarles una ecualización del histograma y posteriormente el algoritmo de retina.	52
Figura 5.13 Imágenes con iluminación controlada e iluminación variable, resultantes de aplicarles una ecualización del histograma y posteriormente el algoritmo de retinex multi-escala.	52
Figura 5.14 Gráfica comparativa de los métodos ecualización de histograma con retina y ecualización de histograma con retinex multi-escala, respecto a la métrica SSIM.	53
Figura 5.15 Imágenes con iluminación controlada e iluminación variable, resultantes de aplicarles una ecualización adaptativa y posteriormente el algoritmo de retina.	53
Figura 5.16 Imágenes con iluminación controlada e iluminación variable, resultantes de aplicarles una ecualización adaptativa y posteriormente el algoritmo de retinex multi-escala.	54
Figura 5.17 Gráfica comparativa de los métodos ecualización adaptativa con retina y ecualización adaptativa con retinex multi-escala, respecto a la métrica SSIM.....	54
Figura 5.18 Grafica comparativa de los algoritmos de pre-procesamiento respecto a la métrica SSIM.....	55
Figura 5.19 Muestra de la imagen de entrenamiento y las correspondientes a la etapa de evaluación del banco de imágenes FEI Face Database.	56
Figura 5.20 Comparativa de exactitud de los algoritmos de pre-procesamiento respecto al banco de imágenes FEI Face Database.....	57
Figura 5.21 Muestra de la imagen de entrenamiento y las correspondientes a la etapa de evaluación del banco de imágenes Faces94.	58
Figura 5.22 Comparativa de exactitud de los algoritmos de pre-procesamiento respecto al banco de imágenes Faces94.	59
Figura 5.23 Muestra de la imagen de entrenamiento y las correspondientes a la etapa de evaluación del banco de imágenes propio.	59
Figura 5.24 Comparativa de exactitud de los algoritmos de pre-procesamiento respecto a al banco de imágenes propio.	61
Figura 5.25 Comparación de resultados de clasificación respecto a los bancos de imágenes.	62
Figura 5.26 Grafica de tiempos de entrenamiento expresados en segundos.	63
Figura 5.27 Grafica de tiempos de monitoreo promedio por imagen expresados en segundos.	64
Figura A.1 Pantalla “Descargas de Visual Studio”.	70
Figura A.2 Pantalla de configuracion del sistema.....	71
Figura A.3 Pantalla de propiedades del sistema.....	71
Figura A.4 Ventana de edición de variables del sistema.	72

Figura A.5 Selección de carpetas donde está contenida la carpeta OpenCV.	72
Figura A.6 Selección de la versión de visual studio que se desea configurar.	72
Figura A.7 Proceso de configuración.	73
Figura A.8 Selección de los módulos adicionales.	74
Figura A.9 Generación de la configuración.	74
Figura A.10 Pantalla de selección del archivo .sln.	75
Figura A.11 Pantalla de configuración para Dlib.	75
Figura A.12 Selección del archivo de configuración .sln para Dlib.	76
Figura A.13 Configurar visual studio en sección VC++	76
Figura A.14 Configurar visual studio en sección directorio de archivos de inclusión. ..	76
Figura A.15 Configurar visual studio en sección directorios de archivos de biblioteca. ..	77
Figura A.16 Configurar visual studio en sección directorio de bibliotecas adicionales. ..	77
Figura A.17 Configurar visual studio en sección dependencias adicionales.	78
Figura A.18 Configurar visual studio en sección directorios de inclusión adicionales. ..	79
Figura A.19 Configurar visual studio en sección directorios de inclusión adicionales para Qt.	79
Figura A.20 Ventana de edición de variables del sistema.	79
Figura B.1 Constancia por la publicación del artículo titulado “Comparación de algoritmos de clasificación con reducción de variables para reconocimiento facial” en el congreso CIINDET 2018.	80
Figura B.2 Constancia por la participación como congresista con el artículo titulado “Comparación de algoritmos de clasificación con reducción de variables para reconocimiento facial” en el congreso CIINDET 2018.	80
Figura B.3 Reconocimiento por la participación como ponente de la conferencia “Reconocimiento automático del rostro para verificación de identidad para evaluación en línea” en el marco del evento: Escuela de inteligencia artificial y robótica 2018, llevado a cabo en las instalaciones de la universidad tecnológica Emiliano Zapata.	81
Figura B.4 Constancia por la participación con el póster “Reconocimiento automático del rostro para verificación de identidad para evaluación en línea”, en la 1A jornada de Ciencia y Tecnología aplicada.	81
Figura B.5 Reconocimiento por la participación como concursante en la categoría innovación en el Advanced Robotics and Drone Competition que se celebró en el Congreso Internacional de Ingeniería Mecatrónica, Electrónica y Automática (ICMEAE).	82
Figura B.6 Constancia por la participación con el poster titulado “Suplantación de identidad en imágenes faciales” en la Segunda Jornada de Ciencia y Tecnología aplicada.	82
Figura B.7 Reconocimiento por la participación como ponente de la conferencia denominada “Reconocimiento automático del rostro para verificación de identidad para evaluación en línea” en el marco del 1er coloquio “Artificial Intelligence” realizada en el Instituto Tecnológico de Iguala (ITI).	83

Lista de Tablas

Tabla 2.1 Resumen de artículos del estado del arte acerca del reconocimiento facial. .11	11
Tabla 2.2 Resumen de artículos del estado del arte con iluminación variable.16	16
Tabla 2.3 Resumen de sistemas del estado de la práctica.21	21
Tabla 2.4 Sistemas del estado de la práctica de los cuales no se encuentra información disponible.....21	21
Tabla 5.1 Resultados de detección del rostro.....44	44
Tabla 5.2 Imágenes de localización del rostro con diferentes ángulos.46	46
Tabla 5.3 Resultados de clasificación del banco de imágenes FEI Face Database en combinación con el algoritmo de pre-procesamiento propuesto FRRM.....56	56
Tabla 5.4 Resultados de exactitud del banco de imágenes FEI Face Database respecto a los algoritmos de pre-procesamiento.57	57
Tabla 5.5 Resultados de clasificación del banco de imágenes Faces94 en combinación con el algoritmo de pre-procesamiento propuesto FRRM.58	58
Tabla 5.6 Resultados de exactitud del banco de imágenes Faces94 respecto a los algoritmos de pre-procesamiento.....58	58
Tabla 5.7 Resultados de clasificación en combinación con el algoritmo propuesto FRRM respecto al banco de imágenes propio, expresado en porcentajes.....60	60
Tabla 5.8 Resultados de exactitud del banco de imágenes propio respecto a los algoritmos de pre-procesamiento.....60	60
Tabla 5.9 Resultados de los tiempos de entrenamiento expresados en segundos.62	62
Tabla 5.10 Resultados de los tiempos de monitoreo expresados en segundos.63	63
Tabla 6.1 Avance de logros por objetivo específico.67	67

Glosario

3D	3 Dimensiones
WFND	<i>Weighted Fractal Neighbor Distance</i> (Distancia ponderada del fractal vecino)
ICA	<i>Independent component analysis</i> (Análisis de componentes independientes)
Dlib	Biblioteca de software multiplataforma de propósito general escrita en el lenguaje de programación C ++
FRRM	Capa de foto receptores con retinex multi-escala
CS-LBP	<i>Center symmetric local binary patterns</i> (patrones binarios locales simétricos)
CLAHE	<i>Contrast Limited Adaptive Histogram Equalization</i> (Ecuación adaptativa del histograma)
CNN	<i>Convolutional neural network</i> (Red neuronal convolucional)
DoG	Diferencias gaussianas
DCT	<i>Discrete cosine transform</i> (Transformada de coseno discreta)
DMSC	<i>Discriminative multi-scale sparse coding</i> (Modelo discriminativo de codificación dispersa a múltiples escalas)
DCP	<i>Dual-cross patterns</i> (Patrones de doble cruz)
Qt	Entorno de trabajo multi-plataforma orientado a objetos
FN	Falsos negativos
FP	Falsos positivos
FND	<i>Fractal Neighbor Distance</i> (Distancia del fractal vecino)
FNDC	<i>Fractal Neighbor Distance based Classification</i> (Distancia del fractal vecino basada en clasificación)
GF	<i>Gabor energy filters</i> (Filtros de energía Gabor)
GOLD	<i>Gaussian of Local Descriptors</i> (Descriptores locales gaussianos)
HASC	<i>Heterogeneous Auto-Similarities of Characteristics</i> (Auto-Similitudes de Características Heterogéneas)
EH	<i>Histogram Equalization</i> (Ecuación del histograma)
SQI	<i>Imagen Self-quotient</i>
IPL	<i>In-ner plexiform layer</i> (Capa plexiforme interna)
IFE	Instituto federal electoral
INE	Instituto nacional electoral
LMA	<i>Largest matching area</i> (Área de mayor coincidencia)

OpenCV	Librería de visión artificial
LDA	<i>Linear Discriminant Analysis</i> (Análisis Lineal Discriminante)
MDML-DCP	<i>Multi-Directional Multi-Level Dual-Cross Patterns</i> (Patrones de doble cruz multinivel multidireccional)
MR	<i>Multi-scale Retinex</i> (Retinex multiescala)
MLP	<i>Multi-Layer Perceptron</i> (Perceptron multi capa)
OPL	<i>Outer plexiform layer</i> (Capa plexiforme externa)
LBP	<i>Local binary patterns</i> (Patrón binario local)
PCA	<i>Principal component analysis</i> (Análisis de componente principal)
RGB	Rojo, Verde, Azul
SSR	<i>Single scale Retinex</i> (Retinex de una sola escala)
SILT	<i>Sparse illumination transfer</i> (Transferencia de iluminación dispersa)
SR	<i>Sparse representation</i> (Representación dispersa)
SRC	<i>Sparse representation-based classificatio</i> (Representación dispersa basada en clasificacion)
SSIM	<i>Structural Similarity Index</i> (Índice de similitud estructural)
SVM	<i>Support vector machines</i> (Maquinas de vector soporte)
TT	<i>Tan and Triggs</i>
SSPP	Una sola muestra por persona
VN	Verdaderos negativos
VP	Verdaderos positivos
VIP	<i>Very important person</i> (Persona muy importante)
WPD	<i>Wavelet Packet Decomposition</i> (Descomposición del paquete wavelet)
YCbCr	Y: luminancia; Cb, Cr: características colorimétricas del color
ZM	<i>Zernike moments</i> (Momentos de Zernike)

CAPÍTULO 1

Introducción

A pesar de los avances producidos en el campo del *e-learning*, existe un número muy reducido de trabajos que permiten, mediante un mecanismo adecuado, la verificación de identidad de aspirantes cuando realiza un examen en línea, con el objetivo de disminuir el número de engaños que se producen en estos sistemas.

Si bien es cierto que existen sistemas de exámenes en línea que incluyen herramientas para la autenticación de los aspirantes, estas herramientas sólo verifican la identidad del usuario en el momento de inicio de sesión, normalmente, a través de un nombre de usuario y una contraseña, pero esta autenticación no garantiza que la persona que ha iniciado la sesión sea quien dice ser. Ante este problema, el uso de un software de verificación de identidad en el transcurso del examen en línea puede permitir identificar y prevenir engañar al sistema.

Esta tesis propone y desarrolla un sistema que permite garantizar que el aspirante sea realmente quien dice ser por medio de un sistema de verificación de identidad.

En este sentido, existen diferentes tecnologías para verificar la identidad de los aspirantes, y uno de esos métodos biométricos que está ocupando un gran peso en los últimos años es el reconocimiento facial [1]; sin embargo, a pesar de que el reconocimiento frontal de la cara ya ha alcanzado una alta precisión en bases de datos disponibles internacionalmente, el reconocimiento facial en entornos hostiles como lo son los ambientes naturales sigue siendo un problema desafiante.

1.1 Planteamiento del problema

El reconocimiento facial en ambientes naturales ha sido y sigue siendo una tarea desafiante para la verificación de identidad al tener que considerar factores ambientales cambiantes tales como, iluminación deficiente, diferentes orientaciones, oclusión parcial del rostro, además de imágenes de baja resolución.

El problema es realizar la identificación cíclica de los aspirantes que presentan el examen de admisión en línea, ya que es importante verificar que la persona que realiza el examen es la misma que se registró en el sistema (imágenes capturadas al inicio del examen) mostrando su identificación (IFE o INE). Este proceso de reconocimiento facial debe contar con cierta robustez ante los aspectos antes mencionados.

1.2 Complejidad del problema

La complejidad del problema se ve reflejada en los siguientes puntos:

- Realizar el reconocimiento facial con condiciones de iluminación variantes.
- Considerar fondos que pueden ser estáticos y dinámicos; es decir, con un fondo en movimiento.
- Analizar imágenes con diferente nivel de resolución.
- Trabajar con rostros con diferente pose.
- Realizar la verificación de identidad en un tiempo cercano al real.

1.3 Objetivo general

Desarrollar un sistema de visión artificial que realice el reconocimiento facial de una persona, de manera automática, para la verificación de la identidad de aspirantes que presentan examen de admisión en línea.

1.3.1 Objetivos específicos

- Revisar en el estado del arte las técnicas utilizadas para reconocimiento facial con limitaciones de iluminación.
- Analizar los componentes del rostro que describen mejor a una persona en un ambiente natural.
- Diseñar e implementar un sistema que realice el reconocimiento con diferentes condiciones de intensidad luminosa.
- Identificar a la persona en un ambiente no controlado.
- Proporcionar una respuesta cercana al tiempo real.

1.4 Alcances y limitaciones

Alcances

- El sistema de reconocimiento de patrones debe llevar a cabo la localización de puntos característicos del rostro para la descripción a través de distancias faciales; ante, diferentes condiciones de iluminación.
- El sistema trabaja con fondos complejos y en algunos casos dinámicos (con movimiento).
- La adquisición es a través de cámaras web, con diferentes resoluciones.
- La respuesta del sistema se proporciona en un máximo de 30 segundos.
- La imagen analizada debe contener el rostro de una persona.
- La verificación de los aspirantes durante la realización del examen se lleva a cabo mediante la comparación de las imágenes adquiridas en el transcurso de esta actividad con las imágenes adquiridas, previamente, durante su registro e identificación en el sistema.

Limitaciones

- No se garantiza la correcta verificación de identidad si los aspirantes se cubren el rostro o utilizan accesorios como gorras y lentes oscuros.
- No se contempla la integración con el sistema de exámenes de admisión.
- La imagen adquirida por la cámara web debe contener el rostro de al menos una persona.
- No se compara la imagen de la INE con las capturas de la cámara web.

1.5 Metodología de solución

En esta sección se describe la propuesta de solución para resolver el objetivo planteado en este tema de tesis. La metodología se divide en seis fases como se ilustra en la figura 1.1.

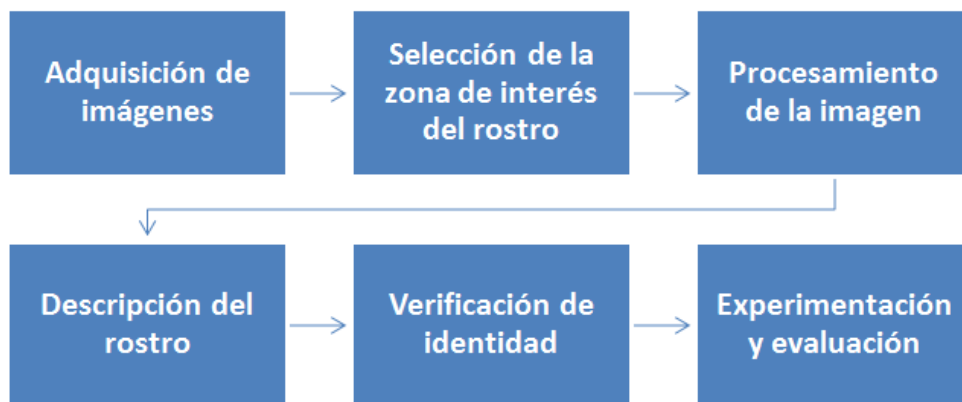


Figura 1.1 Esquema de metodología de solución.

A continuación se describe brevemente el objetivo de cada fase del sistema. En el capítulo 4 se detalla la implementación de los mismos.

1) Adquisición de imágenes

Este apartado describe la adquisición de las imágenes la cual se integra de 2 etapas. Para el entrenamiento se utiliza la imagen enviada, vía correo electrónico, para el registro de cada aspirante. El monitoreo de aspirantes se realiza en línea (etapa de evaluación) a través la captura de su rostro, vía cámara web. Para las pruebas se realizó mediante videos previamente grabados.

2) Selección de la zona de interés del rostro

Este apartado describe el proceso para seleccionar la zona de interés, que en este caso es un rostro de frente, descartando toda la demás información presente en la imagen.

3) Procesamiento de la imagen

Este apartado se encarga de procesar la imagen para disminuir los efectos de una iluminación variante, propia de un ambiente no controlado y, así poder tener una imagen con mejor calidad para las siguientes fases.

4) Descripción del rostro

Este apartado se encarga de localizar puntos característicos en el rostro y realizar la descripción del rostro en términos de distancias faciales.

5) Verificación de identidad

Como su nombre lo dice, este apartado tiene el objetivo de llevar a cabo la identificación del aspirante utilizando el modelo de clasificación generado en el entrenamiento.

6) Experimentación y evaluación

Este apartado consiste en diseñar y realizar la experimentación del sistema en los módulos de pre-procesamiento de las imágenes como en clasificación, reportando y analizando los resultados a través de diferentes medidas de evaluación.

1.6 Organización de la tesis

Capítulo 1: Contiene el análisis del problema de este trabajo.

Capítulo 2: Revisa el estado del arte y la discusión correspondiente.

Capítulo 3: Muestra el marco teórico.

Capítulo 4: Este capítulo comprende el diseño e implementación de la solución para el problema de tesis.

Capítulo 5: Describe el objetivo, el diseño y los resultados de las pruebas que se realizaron para cada módulo del desarrollo de la tesis.

Capítulo 6: Contiene las conclusiones generales, la discusión sobre el tema de tesis, trabajos futuros y aportaciones.

Anexos. Contienen elementos importantes para el desarrollo de esta tesis, los cuales son: Manual de instalación (OpenCV, Dlib y Qt).

CAPÍTULO 2

Estado del arte

En los últimos años se han desarrollado diversos trabajos enfocados al reconocimiento facial con problemas como lo son: pose ante una cámara, oclusión parcial del rostro, expresiones faciales, iluminación no controlada, etc.

En este capítulo se describen técnicas utilizadas en la actualidad para dar solución a alguno o varios de los problemas anteriormente mencionados. El estado del arte está dividido en dos secciones que son: Sistemas de reconocimiento facial en general y artículos que tienen como objetivo el ser robustos ante variaciones de iluminación en el reconocimiento facial, así como una recopilación de sistemas que están actualmente en el mercado, los cuales utilizan el reconocimiento facial (estado de la práctica).

2.1 Reconocimiento facial

Existen diferentes tecnologías para realizar reconocimiento facial, a continuación se muestra un análisis del estado del arte sobre los métodos y técnicas que se han empleado en los últimos años.

En el artículo titulado *reconocimiento del rostro robusto a la pose con Huffman-LBP utilizando la estrategia de divide y vencerás* [2], se desarrolló una extracción de características tipo patrón binario local (LBP), que modifica la regla de código de Huffman, el cual lleva por nombre Huffman-LBP para extraer rasgos faciales. Además, se aplica una estrategia de división y regla a la representación de la cara y la clasificación, que tiene como objetivo mejorar el rendimiento del reconocimiento en toda pose. En otra investigación titulada *reconocimiento rápido de rostros basado en la teoría fractal* [3] también abordan el problema del reconocimiento facial mediante la división del rostro, se introduce un método rápido de codificación fractal para el reconocimiento facial denominado clasificación de distancia del vecino fractal (FNDC), que mejora los métodos tradicionales de reconocimiento de fractales como son la distancia del vecino del fractal ponderado (WFND) y la distancia del fractal vecino (FND), utilizando información de clase para establecer umbrales entre clases y dentro de clases para acelerar la velocidad de reconocimiento.

En la investigación titulada *codificación dispersa multi-escala discriminativa para reconocimiento de rostros con oclusión* [4] abordan el problema del reconocimiento facial partiendo de una sola muestra por persona (SSPP), es un problema y desafío importante para los sistemas de reconocimiento facial, debido a la falta de información de datos de muestra. En el trabajo se comenta que existen métodos para superar las variaciones en iluminación, expresión y pose, que abordan el problema de SSPP. Sin embargo, no son robustos cuando las muestras de prueba son con diferentes tipos de oclusión (Figura 2.1).

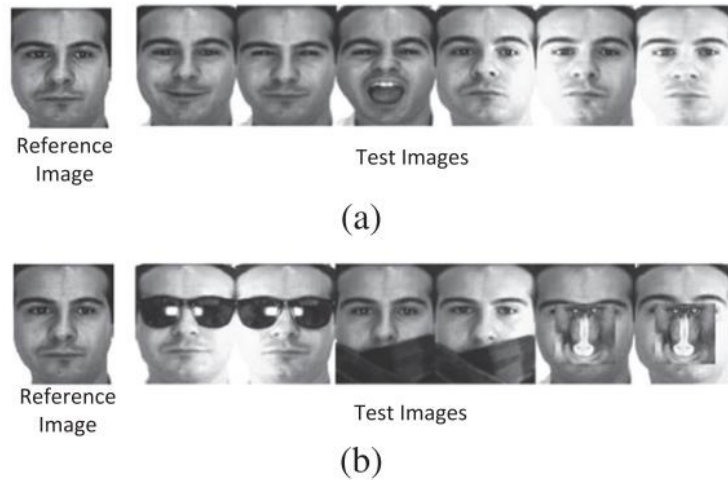


Figura 2.1 (a) Imágenes con cambios de expresión e iluminación. (b) imágenes con diferentes tipos de oclusiones (disfraz real y oclusión de bloques aleatorios). La tarea de reconocimiento en estas imágenes (b) es mucho más difícil que la tarea de reconocimiento que en (a) [4]

Este trabajo, propone un modelo discriminativo de codificación dispersa a múltiples escalas (DMSC) para abordar el problema de una sola muestra por persona. A continuación se muestra la metodología de solución para el reconocimiento de rostros con oclusión:

- i) Se aprende un diccionario auxiliar a partir de datos externos basados en PCA para modelar las posibles variaciones de oclusión.
- ii) Se desarrolló una medida de error a múltiples escalas para detectar y descartar píxeles atípicos debido a la oclusión, lo que es eficaz para mejorar la dispersión, la solidez y la capacidad de discriminación.
- iii) Se usa el algoritmo DMSC basado en la combinación del diccionario aprendido y la estrategia de medición de errores a múltiples escalas.

En el trabajo, *Reconocimiento facial: desafíos, logros y direcciones futuras* [5] el objetivo es analizar los desafíos importantes que conlleva la adaptación de los algoritmos de reconocimiento facial existentes para construir sistemas exitosos que puedan emplearse en el mundo real, los desafíos que se analizan son variaciones de iluminación, variaciones de postura, variaciones por envejecimiento, expresión facial / estilo facial y oclusión. Los enfoques tradicionales que se analizaron, los dividieron en dos categorías las cuales se muestran en la figura 2.2.

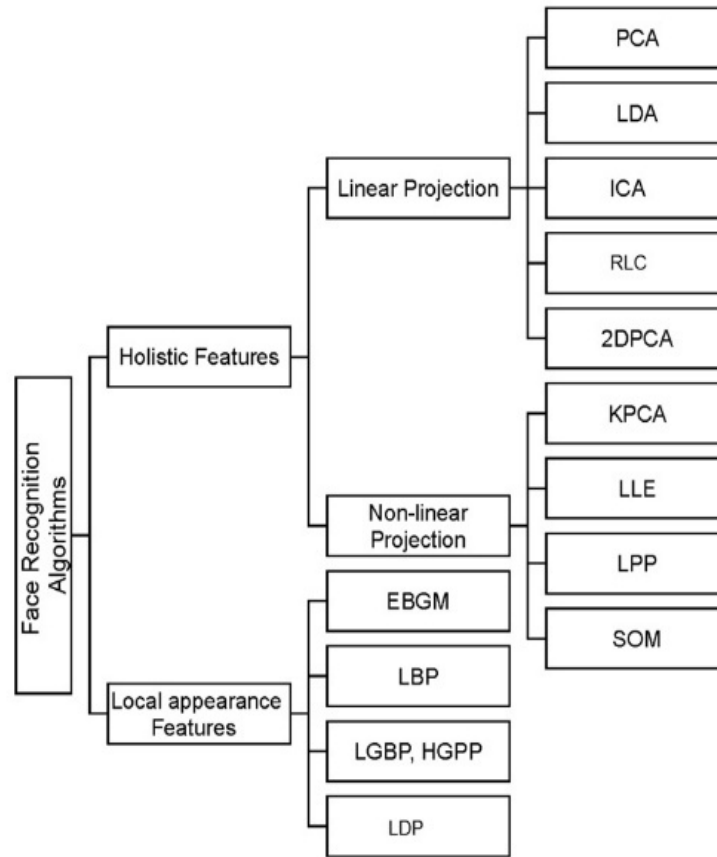


Figura 2.2 Categorías de algoritmos de reconocimiento de rostros tradicionales. [5]

En el estudio sobre técnicas de reconocimiento facial [6], al igual que en el anterior artículo, se analizaron diferentes métodos para el reconocimiento facial; sin embargo, en este estudio se propone una división en tres partes, que son:

- Métodos holísticos de coincidencia, en los cuales toda la región de la cara se tiene en cuenta como datos de entrada en el sistema de captura de rostros, un ejemplo es el análisis de componente principal (PCA). Esta técnica es utilizada en el trabajo; *Sistema de asistencia basado en reconocimiento facial usando Algoritmos PCA* [7], el cual consiste en registrar la asistencia de alumnos a clase, enfrentando los problemas de postura y variación de iluminación.
- Métodos estructurales, se localizan y extraen en primer lugar las características esenciales, como la ubicación de los ojos, la nariz y la boca; y sus estadísticas nativas (geométricas o de apariencia) se incorporan a un clasificador estructural.
- Métodos híbridos, estos utilizan una combinación de métodos de extracción holística y de características locales, comúnmente se usa tecnología 3D.

En la investigación, *Patrones binarios locales extendidos para el reconocimiento facial* [8], si bien no están enfocados en una mala iluminación, utilizan LBP basado en textura para lograr el reconocimiento facial. De la textura se extraen características de

intensidad local, características locales basadas en la diferencia angular y características locales basadas en la diferencia radial, para obtener el vector de características. Para poder reducir la dimensionalidad usan PCA (Análisis de componentes principales) y así producir características compactas, robustas y discriminantes, que ayuden al reconocimiento facial.

Dentro de las técnicas recientes aplicadas al reconocimiento facial se encuentra el aprendizaje profundo utilizado en el artículo: *Análisis exhaustivo del aprendizaje profundo basado en la representación para el reconocimiento facial* [9], el cual hace uso de los modelos *VGG-Face* y *Lightening CNN*. La investigación se centra en utilizar el aprendizaje profundo sobre las imágenes sin utilizar algún preprocesamiento, la desventaja es que utilizan grandes conjuntos de imágenes para el entrenamiento.

En la tabla 2.1 se muestra un resumen sobre los artículos del estado de arte acerca del reconocimiento facial.

Tabla 2.1 Resumen de artículos del estado del arte acerca del reconocimiento facial.

Artículo	Objetivo	Aportación	Técnicas
Reconocimiento del rostro robusto a la pose con Huffman-LBP utilizando la estrategia de divide y vencerás [2]	Realizar un sistema robusto a la pose de la cara mediante la unión de dos algoritmos Huffman y LBP para la extracción de características.	Técnicas para la extracción de características.	Huffman LBP
Reconocimiento rápido de rostros basado en la teoría fractal [3]	Realizar un reconocimiento facial para aminorar los efectos de la iluminación, pose y oclusión	Técnicas de reconocimiento facial para abordar la pose.	Distancia del fractal vecino, Clasificación de distancia basada en el fractal vecino.
Codificación dispersa multi-escala discriminativa para reconocimiento de rostros con oclusión [4]	Reconocimiento facial con diferentes tipos de oclusión basado en una sola muestra por persona.	Técnicas de reconocimiento facial con oclusión	Modelo discriminativo de codificación dispersa a múltiples escalas (DMSC).
Reconocimiento facial: desafíos, logros y direcciones futuras [5]	Analizar los desafíos importantes que conlleva la adaptación de los algoritmos de reconocimiento facial existentes para construir sistemas exitosos que puedan emplearse en el mundo real.	Técnicas utilizadas para abordar variaciones de iluminación, variaciones de postura, variaciones por envejecimiento, expresión facial / estilo facial y oclusión.	<ul style="list-style-type: none"> • LBP • Proyección no-lineal • PCA • LDA • ICA
Estudio sobre técnicas de reconocimiento facial [6]	Análisis de los componentes de un sistema de reconocimiento facial así como métodos para abordar el problema.	Problemas que enfrentan el reconocimiento facial y métodos que son usados para abordar el problema.	<ul style="list-style-type: none"> • PCA • Métodos de emparejamiento holístico • Métodos estructurales • Métodos híbridos
Sistema de asistencia basado en reconocimiento facial usando Algoritmos PCA [7]	Sistema para llevar a cabo el registro de asistencia de alumnos dentro del salón de clase por medio de reconocimiento facial.	Como se aborda el problema del reconocimiento facial en un ambiente natural con variación de pose.	<ul style="list-style-type: none"> • PCA • ICA • MLP • SVM
Patrones binarios locales extendidos para el reconocimiento facial [8]	Reconocimiento facial basado en texturas LBP.	Extracción de características basadas en LBP.	<ul style="list-style-type: none"> • LBP • PCA
Análisis exhaustivo del aprendizaje profundo basado en la representación para el reconocimiento facial [9]	Reconocimiento facial basado en aprendizaje profundo	Realizar reconocimiento facial sin utilizar el preprocesamiento.	<ul style="list-style-type: none"> • Modelo VGG-Face • Lightning CNN

2.2 Reconocimiento facial con iluminación variable

En las últimas décadas, el reconocimiento facial se ha convertido en el centro de investigación con particular interés en los ambientes no controlados, para dar solución a problemas como una iluminación deficiente, oclusión parcial o total, escasas muestras de entrenamiento y expresiones faciales, entre otras. Han surgido diferentes métodos y técnicas para abordar el reconocimiento facial con una o varias limitaciones. A continuación se muestra un análisis del estado del arte sobre los métodos y técnicas que se han empleado en los últimos años.

En un estudio del 2015 se realizó una *comparativa de algoritmos para abordar el problema de la iluminación en el reconocimiento facial bidimensional* [10]. Los algoritmos que mencionan son: Retinex de una sola escala (SSR), Retinex multiescala (MR), Imagen Self-quotient (SQI), Transformada de coseno discreta (DCT) basada en normalización, Transformada Wavelet basada en normalización, normalización *Tan and Triggs* (TT), *Weberfaces* y modelo de retina. En la figura 2.3 se muestra una comparación visual de la aplicación de los algoritmos mencionados.

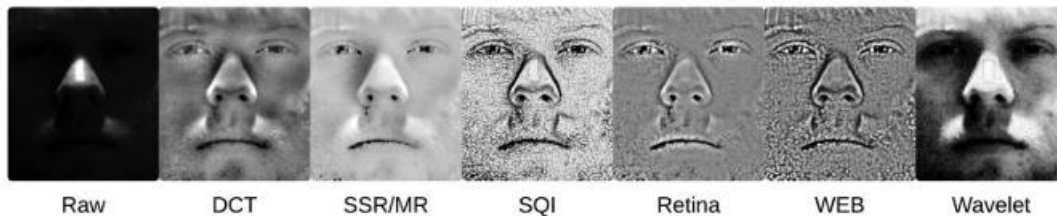


Figura 2.3 Comparación visual de imágenes faciales de la base de datos Yale face database B, después de aplicar diferentes métodos [10].

Los autores proponen la creación de dos categorías: iluminar y quitar iluminación; los métodos de iluminación intentan coincidir con las condiciones de iluminación, mientras que los métodos de quitar iluminación intentan suprimir las variaciones. Dentro de los métodos de iluminación mencionan: basado en Bootstrap, basado en diccionario; mientras que en los métodos de quitar iluminación se mencionan: basado en fotometría, transformada discreta de coseno, descomposición Wavelet, basada en características, patrón binario local (LBP) y Sub-espacios lineales. Los resultados en su investigación indican que los métodos de iluminación son una mejor y práctica solución para abordar los desafíos de iluminación si el objetivo es la solidez a diferentes efectos, como la iluminación u oclusión. Ya que se adaptan fácilmente a este propósito sin disminuir el rendimiento del reconocimiento general.

Una investigación que ocupa la descomposición de wavelet es, un *método nuevo de reconocimiento facial invariante a iluminación basado en PCA y WPD usando el espacio de color YCbCr* [11], el cual a diferencia de las investigaciones pasadas, no ocupa imágenes en escala de grises, si no que convierte la imagen de la cara en componentes Y, Cb y Cr. Cada componente se divide con la descomposición de paquetes Wavelet y posteriormente se reduce la dimensionalidad con PCA para la clasificación.

En otra investigación titulada, *Evaluación de métodos de reconocimiento facial en ambientes sin restricciones* [12], además de comparar técnicas para abordar el problema de la iluminación, realizan una comparación para revisar que tan robustos son respecto a expresiones faciales y postura, además de analizar sus pro y contras. Los autores proponen una clasificación en dos grupos: Métodos basados en la apariencia y métodos basados en modelos. Los métodos basados en apariencia utilizan características de textura holística que se aplican a regiones de toda la cara o específicas en una imagen de la cara, mientras que los basados en modelos emplean la forma y la textura de la cara.

Una investigación que aborda un desafío diferente a los antes mencionados es, *Aprendizaje y transferencia de iluminación dispersa para el reconocimiento facial de una sola muestra con corrupción y desalineación de la imagen* [13], en la que se utiliza la técnica de aprendizaje y transferencia de iluminación dispersa (SILT), esta se basa en la representación dispersa (SRC) para la alineación y el reconocimiento de muestras individuales. El problema de este enfoque es que se necesitan imágenes auxiliares con diferentes iluminaciones para poder realizar el reconocimiento facial. Otra investigación que trabaja con una sola muestra es, *Áreas de mayor coincidencia para un sistema de reconocimiento facial robusto a iluminación y oclusión* [14], en el que se compara la similitud de imágenes de rostros, realizando una búsqueda de la mayor área de coincidencia (LMA) en cada ubicación de cara. En contraste con los métodos existentes de normalización de iluminación y modelos de oclusión basados en áreas locales de tamaño fijo de la imagen, la comparación de imágenes basada en las LMA's mejora la discriminación entre diferentes personas y, al mismo tiempo, optimiza la precisión de iluminación.

En el artículo de revisión: *Reconocimiento facial mediante técnicas de iluminación* [15], se describen diferentes técnicas para abordar el problema de las variaciones de iluminación en el reconocimiento facial, con la finalidad de poder ser base para investigaciones futuras que pretendan abordar dicho problema. La investigación se enfoca en la normalización de la iluminación antes de la etapa del procesamiento de la imagen, debido a que, lo consideran la mejor manera de abordar el problema. Los autores se enfocan, para realizar el análisis, en seis técnicas que son: Transformación Coseno Discreta, Wavelet, normalización basada en filtrado orientable, diferencia gaussiana y difusión anisotrópica modificada para la normalización.

Otros artículos que utilizan la normalización de iluminación son: *Técnica de reconocimiento facial robusta a iluminación variable* [16] y *Sistema efectivo de Reconocimiento facial para iluminación no controlada* [17]. En este último, la imagen de entrada se divide en dos componentes: un componente de alta frecuencia y uno de baja frecuencia, con el fin de obtener un suavizado de la imagen y un realce de los contornos, que posteriormente fusionan con la imagen original, con el fin de dar realce a las características del rostro.

El artículo *conjunto de descriptores de textura y clasificadores para reconocimiento facial* [18], se basa en la textura como la mejor manera de abordar el problema del reconocimiento facial en ambientes naturales: la entrada es una imagen frontal, y la

salida es un conjunto de imágenes pre-procesadas de acuerdo con los siguientes enfoques: Retinex adaptable de escala única, Suavizado anisotrópica, Diferencia de gaussianas, un enfoque basado en la transformada de coseno discreta de baja frecuencia, la ecualización de histograma local orientada, Retinex multi-escala, la normalización de suavizado isotrópico y las caras de gradiente. En la investigación, *reconocimiento facial mediante técnicas de iluminación* [15], también utilizan las técnicas de ecualización del histograma local y normalización de suavizado isotrópico, para el preprocesamiento. Para la extracción de características emplean Patrones Binarios Locales (LBP), Histograma de Gradientes, Auto-Similitudes de Características Heterogéneas (HASC), descriptores locales gaussianos (GOLD), y codificación binaria monogénica. Posteriormente para la reducción de dimensionalidad usan ACP.

LBP es de las técnicas más utilizadas dentro de la literatura para abordar el problema de la iluminación. En el artículo, *Codificación y descodificación de patrones binarios locales para la normalización de iluminación del rostro* [19], utilizan esta técnica, ya que mencionan que LBP conserva la tolerancia a los cambios de iluminación y utilizan el mapeo inverso desde el dominio LBP al dominio de píxeles para recuperar la iluminación de la cara normalizada con alta fidelidad. Sin embargo, dicha técnica no sólo es empleada para resolver el problema de iluminación, en el artículo, *Reconocimiento facial considerando desenfoque no uniforme de movimiento, iluminación y pose* [20], es empleado para abordar el problema de desenfoque causado por movimiento y cara borrosa.

Una investigación que enfatiza el tema de baja resolución es, *Reconocimiento facial en baja resolución considerando variaciones en postura e iluminación* [21] el cual trabaja con una imagen de entrenamiento de alta resolución y las imágenes adquiridas en baja resolución; el objetivo es nivelar la resolución de las dos imágenes para lograr el reconocimiento facial. Otra investigación que trabaja con postura e iluminación es, *Reconocimiento facial con variaciones de pose e iluminación a través de una representación dispersa y diccionario de iluminación* [22] a diferencia del anterior trabajo utilizan la representación dispersa (SR) y diferentes muestras de poses para calcular la similitud entre las imágenes.

En la investigación *Reconocimiento facial usando el patrón binario local aumentado y la métrica de disimilitud de Bray Curtis* [23], abordan el problema del reconocimiento facial en entornos no controlados, su método basado en textura, consiste en reemplazar patrones no uniformes con el valor mayoritario de patrones uniformes y combinarlos con patrones vecinos uniformes para extraer información con respecto a los descriptores locales.

En la investigación, *Reconocimiento facial invariante a expresiones usando patrones binarios locales y transformada de contorno* [24], abordan el problema de la iluminación y oclusión parcial enfocándose sólo en las variaciones de las imágenes, utilizan LBP de bloques múltiples para extraer las variaciones locales densas en el dominio espacial y luego la imagen original se somete a la transformada de contorno para obtener sub-bandas. Específicamente, la transformada de contorno muestra propiedades como direccionalidad y anisotropía, por lo tanto, da como resultado la

extracción de características significativas. En la investigación, *Reconocimiento facial robusto con patrones gradientes binarios estructurales* [25], a diferencia de los anteriores que usan LBP, utiliza una variación llamada, patrones binarios locales simétricos (CS-LBP), para abordar los desafíos de iluminación.

Gabor es una técnica ampliamente utilizada en el reconocimiento facial, en el artículo *Reconocimiento facial con variación de pose con características locales de Gabor mejoradas por la forma activa y modelos estadísticos* [26], se incorporan bloques para compensación de iluminación mediante un método de normalización local y características de Gabor ponderadas por entropía para enfatizar aquellas características que mejoran la identificación. En la investigación, *Reconocimiento facial basado en características Gabor-Zernike* [27] Aplican filtros de energía Gabor (GF) y momentos de Zernike (ZM), donde GF se utiliza para la extracción de características de textura y ZM extrae formas.

En el trabajo de investigación, *Patrones de cruz doble, multiniveles, multidireccionales para un reconocimiento facial robusto* [28], se presenta un esquema para extraer "patrones de doble cruz multinivel multidireccional" (MDML-DCP). Específicamente, el esquema MDML-DCP explota la primera derivada del operador gaussiano para reducir el impacto de las diferencias en la iluminación y luego calcula la función DCP en los niveles holístico y de componentes.

En la tabla 2.2 se muestra un resumen sobre los artículos del estado de arte acerca de la iluminación variable en reconocimiento facial.

Tabla 2.2 Resumen de artículos del estado del arte con iluminación variable.

Artículo	Objetivo	Aportación	Técnicas
Comparativa de algoritmos para abordar el problema de la iluminación en el reconocimiento facial bidimensional [10]	Comparativa de diferentes técnicas de preprocesamiento de imágenes faciales con problemas de iluminación.	Comparación sobre técnicas de preprocesamiento de imágenes faciales.	<ul style="list-style-type: none"> • Retinex de una sola escala (SSR) y multiescala (MR) • Imagen Self-quotient (SQI) • Transformada de coseno discreta (DCT) basada en normalización • Transformada Wavelet basada en normalización • Normalización <i>Tan and Triggs</i> (TT) • <i>Weberfaces</i> • Modelo de retina • Ecuación adaptativa
Un método nuevo de reconocimiento facial invariante a iluminación basado en PCA y WPD usando el espacio de color YCbCr [11]	Reconocimiento facial invariante a la iluminación en imágenes de color.	Trabajar el preprocesamiento de imágenes con problemas de iluminación en color y no en escala de grises.	<ul style="list-style-type: none"> • Wavelets • PCA
Evaluación de métodos de reconocimiento facial en ambientes sin restricciones [12]	Comparación de métodos de reconocimiento facial, invariantes a la iluminación, expresiones faciales y postura.	Técnicas que abordan el problema de reconocimiento facial invariante a expresiones faciales y postura.	<ul style="list-style-type: none"> • LBP • Regresión lineal
Aprendizaje y transferencia de iluminación dispersa para el reconocimiento facial de una sola muestra con corrupción y desalineación de la imagen [13]	Reconocimiento facial con una sola muestra de entrenamiento.	Realizar el reconocimiento facial con un entrenamiento limitado (sólo una muestra) y corrupción de la imagen.	Técnica de aprendizaje y transferencia de iluminación dispersa (SILT).
Áreas de mayor coincidencia para un sistema de reconocimiento facial robusto a iluminación y oclusión [14]	Reconocimiento facial basado en áreas invariantes a iluminación.	Realizar un reconocimiento facial con una sola muestra de entrenamiento sin tomar en cuenta todo el rostro, se toman en cuenta sólo zonas que ayudan al reconocimiento facial con variaciones de iluminación y oclusión parcial.	Áreas de mayor coincidencia (LMA)

Artículo	Objetivo	Aportación	Técnicas
Artículo de revisión: Reconocimiento facial mediante técnicas de iluminación [15]	Análisis de artículos que abordan el problema de reconocimiento facial con variaciones de iluminación.	Revisar las técnicas de preprocesamiento que son utilizadas en investigaciones actuales.	<ul style="list-style-type: none"> • Transformación Coseno Discreta • Wavelet • Normalización basada en filtrado orientable • Diferencia gaussiana • Difusión anisotrópica modificada para la normalización
Técnica de reconocimiento facial robusta a iluminación variable [16]	Reconocimiento facial con variación de iluminación.	Técnicas utilizadas en el preprocesamiento para abordar el problema de la iluminación.	<ul style="list-style-type: none"> • Ecuación del histograma • Retinex multi-escala • Imagen Self-quotient (SQI) • Filtros gaussianos
Sistema efectivo de Reconocimiento facial para iluminación no controlada [17]	Reconocimiento facial basado en frecuencias.	Realce de puntos faciales de interés para mejorar el reconocimiento facial con iluminación no controlada.	<ul style="list-style-type: none"> • Filtros gaussianos
Conjunto de descriptores de textura y clasificadores para reconocimiento facial [18]	Reconocimiento facial basado en texturas.	Realizar el reconocimiento facial con texturas en lugar de usar distancias de puntos de interés.	<ul style="list-style-type: none"> • Retinex adaptable de escala única • Suavizado anisotrópica • Diferencia de gaussianas • Transformada de coseno discreta de baja frecuencia • Ecuación de histograma • Retinex multi-escala • Normalización de suavizado isotrópico
Codificación y descodificación de patrones binarios locales para la normalización de iluminación del rostro [19]	Reconocimiento facial basado en LBP.	Realizar el reconocimiento facial mediante la textura LBP.	LBP
Reconocimiento facial considerando desenfoque no uniforme de movimiento, iluminación y pose [20]	Reconocimiento facial en imágenes con desenfoque.	Utilidad de utilizar textura para el reconocimiento facial en imágenes con desenfoque por movimiento y cara borrosa.	LBP
Reconocimiento facial en baja resolución considerando variaciones en postura e iluminación [21]	Reconocimiento facial en imágenes de baja resolución, con entrenamiento de imágenes de alta resolución.	Realizar el reconocimiento facial en imágenes con variación de pose, iluminación y considerando una adquisición de imágenes con mala resolución.	<ul style="list-style-type: none"> • PCA

Artículo	Objetivo	Aportación	Técnicas
Reconocimiento facial con variaciones de pose e iluminación a través de una representación dispersa y diccionario de iluminación [22]	Reconocimiento facial robusto a pose.	Otra forma de abordar el problema de pose en el reconocimiento facial, basado en diferentes muestras para el cálculo de la similitud entre imágenes.	Representación dispersa (SR)
Reconocimiento facial usando el patrón binario local aumentado y la métrica de disimilitud de Bray Curtis [23]	Reconocimiento facial en entornos no controlados.	Utilizar la textura LBP para el reconocimiento facial en ambientes no controlados.	<ul style="list-style-type: none"> • LBP • Métrica Bray Curtis
Reconocimiento facial invariante a expresiones usando patrones binarios locales y transformada de contorno [24]	Reconocimiento facial invariante a expresiones faciales.	Realizar un reconocimiento facial sólo basado en las variaciones de las imágenes.	<ul style="list-style-type: none"> • LBP
Reconocimiento facial robusto con patrones gradientes binarios estructurales [25]	Reconocimiento facial robusto a iluminación y pose.	Variaciones de LBP para el reconocimiento facial.	<ul style="list-style-type: none"> • Patrones binarios locales simétricos (CS-LBP)
Reconocimiento facial con variación de pose con características locales de Gabor mejoradas por la forma activa y modelos estadísticos [26]	Reconocimiento facial robusto a pose.	Se utiliza Gabor como método de reconocimiento facial basado en texturas en lugar de LBP.	<ul style="list-style-type: none"> • Gabor
Reconocimiento facial basado en características Gabor-Zernike [27]	Reconocimiento facial robusto a iluminación y pose.	Algoritmo basado en texturas de Gabor para el reconocimiento facial.	<ul style="list-style-type: none"> • Gabor • Momentos de Zernike(GM)
Patrones de cruz doble, multiniveles, multidireccionales para un reconocimiento facial robusto [28]	Reconocimiento facial robusto a iluminación.	Reconocimiento facial basado en diferencias de iluminación.	<ul style="list-style-type: none"> • Patrones de doble cruz multinivel multidireccional (MDML-DCP)

2.3 Análisis del estado del arte

Existen diversas técnicas para el preprocesamiento de imágenes con variación de iluminación, pose, expresiones faciales y oclusión parcial. Las técnicas más utilizadas son modelo de retina, retinex y LBP, las cuales son empleadas en diversos artículos del estado del arte para dar solución a más de un problema del reconocimiento facial, teniendo buenos resultados en reconocimiento facial en ambientes naturales. Por ello, se toman como base para el desarrollo de esta tesis.

La extracción de características es una parte importante para el reconocimiento facial, después de un preprocesamiento de la imagen el estado del arte indica que existen dos maneras de obtener el vector de características, por medio de distancias de puntos faciales y por texturas. Los puntos faciales se extraen mediante la distancia euclídea, mientras que en textura la técnica mayormente utilizada es con LBP seguida de Gabor.

2.4 Estado de la práctica

A continuación se describen sistemas de reconocimiento facial que están actualmente en el mercado.

Open Face [29]

Es capaz de captar e identificar los rostros en tiempo real; Necesita al menos diez fotografías de la persona en cuestión para realizar la identificación. Para empezar a utilizarlo se deberá identificar a cada sujeto, completando un apartado informativo; e instalar una serie de requisitos para que el programa aprenda a usar todos sus rasgos como se ilustra en la figura 2.4.

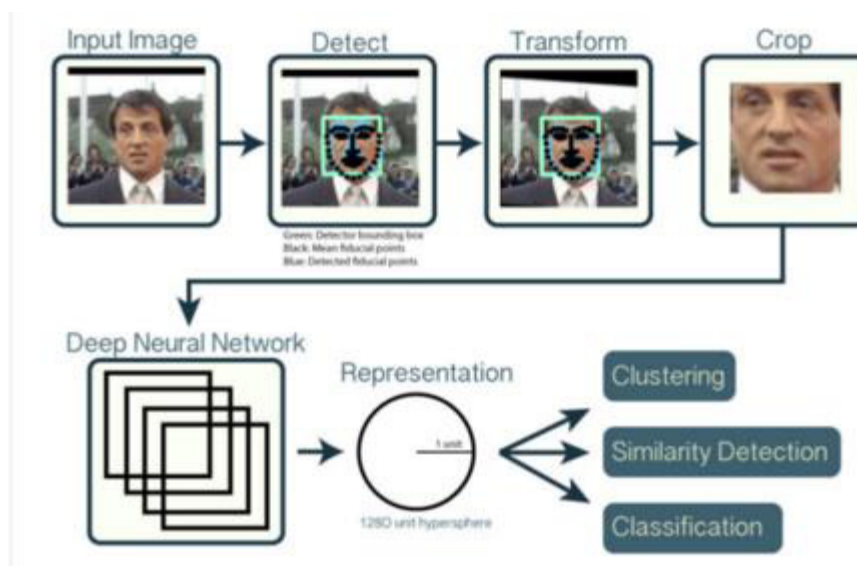


Figura 2.4 Diagrama del proceso de funcionamiento del programa Open Face [29].

Face Net [30]

Sistema que realiza un modelo a partir de las imágenes faciales a un espacio euclidiano compacto donde las distancias corresponden directamente a una medida de similitud de caras. Una vez que se ha producido este espacio, se implementan las tareas tales como reconocimiento facial, verificación y agrupamiento, utilizando técnicas como vectores de características. El método utiliza una red convolucional profunda entrenada para optimizar directamente la propia incrustación, en lugar de una capa de cuello de botella intermedia como en los enfoques previos de aprendizaje profundo.

Axis communications [31]

El software de reconocimiento facial busca las caras que aparecen en el vídeo en tiempo real en una base de datos con imágenes de caras guardadas previamente. Las caras de la base de datos pueden dividirse en diferentes categorías en función de cuál sea su objetivo (control de acceso, detección de personas VIP o identificación de delincuentes conocidos).

BioSurveillance [32]

BioSurveillance, detecta múltiples caras a tiempo real, facilita la gestión de usuarios en cualquier momento y permite entrenar mediante captura de vídeo (*on-the-fly*). Utiliza cámaras de video vigilancia para reconocer las caras de las personas incluidas en la lista negra.

- Ofrece un rendimiento con oclusiones parciales de la cara, gafas, bufandas, gorras, cambios de expresión facial, sombras, altos contrastes y condiciones de iluminación extremas o deficientes, y rotaciones moderadas de la cara.
- Trabaja en movimiento y a distancia, y no requiere la colaboración por parte del usuario a reconocer.
- Las personas pueden inscribirse en el sistema a través de una fotografía, así como mediante el uso de video grabado previamente o en tiempo real.

En la tabla 2.3 se muestra un resumen sobre los sistemas del estado de la práctica acerca del reconocimiento facial, de los cuales se encuentra información disponible como es: si cuenta con acceso al código, licencia, tipo de imágenes que ocupa y lenguaje en que fue programado. En la tabla 2.4 se muestra información sobre sistemas del estado de la práctica de los cuales no se encuentra información disponible, sólo una breve descripción.

Tabla 2.3 Resumen de sistemas del estado de la práctica.

Sistema	Disponibilidad	Licencia	Imágenes	Lenguaje	Comentarios
Open Face	Código abierto	Propiedad de Carnegie Mellon University	Imágenes reales en entorno controlado.	Python	Se requieren 10 capturas de imágenes por persona.
Face Net	No	Google	Imágenes Web de entornos no controlados.	No	Identificación de personas mediante imágenes de la Web.
Axis communications	No	Axis communications	Imágenes reales en entornos no controlados.	No	Resolución mínima de cámara 1080p.
BioSurveillance	No	Herta Security	Imágenes reales en entornos no controlados.	No	Sistema optimizado para arquitecturas GPU, con cámaras de resolución 1080p.

Tabla 2.4 Sistemas del estado de la práctica de los cuales no se encuentra información disponible.

Sistema	Descripción	Dirección
prOctorFREE	<i>ProctorFree</i> ofrece soluciones de autenticación para estudiantes para el entorno de educación en línea. Su tecnología sirve como un puente para verificar que sus estudiantes sean de hecho reales y que sean los mismos estudiantes que se inscribieron en sus cursos en línea.	http://proctorfree.com/student-authentication
eProctoring	Es un servicio que ofrece un entorno seguro para la supervisión en los exámenes en línea con Moodle.	https://eproctoring.com
FindFace	Con sólo una fotografía los usuarios pueden acceder a la información pública de esa persona en el Facebook ruso.	https://findface.ru/

CAPÍTULO 3

Marco teórico

En esta sección se describe el marco teórico para la comprensión de este trabajo. Además, se explica el funcionamiento de los diferentes algoritmos utilizados en el sistema desarrollado, así como las métricas para evaluar los resultados de los algoritmos utilizados para el desarrollo de este trabajo.

3.1 Técnicas de pre-procesamiento

El pre-procesamiento de las imágenes faciales es una etapa fundamental para la detección de los puntos de interés del rostro. En este trabajo se implementaron los algoritmos con un mayor porcentaje de éxito, de acuerdo a la literatura del área: 1) ecualización del histograma, 2) ecualización adaptativa, 3) modelo de retina, 4) retinex multi-escala y 5) se explica el modelo propuesto en esta tesis que es resultado de la fusión de los algoritmos de retina y retinex, para resaltar las zonas de interés del rostro. A continuación se detalla cada una de éstas técnicas.

1) Ecualización del histograma [7] [16] [18] [19]

La ecualización de histogramas es un enfoque popular y poderoso para eliminar la mala iluminación de las imágenes. La función de histograma distribuye el valor de píxel de frecuencia más intensivo. Como consecuencia, se modifica la información del píxel original que cambia el aspecto de la cara como se muestra en la Figura 3.1.

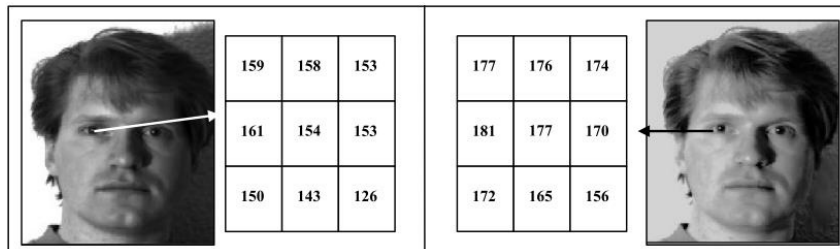


Figura 3.1 Valor de 3x3 píxeles de la porción del ojo izquierdo antes y después de aplicar el histograma [16].

El histograma mejora el contraste de la imagen para que la información que está oculta, debido al factor de iluminación, se haga visible. Se considera que G es la imagen de nivel de gris mientras que i muestra la apariencia de píxeles grises en la imagen n_i veces. Del mismo modo, N es el número total de píxeles en la imagen. Matemáticamente se puede representar como:

$$F(p_k) = \sum_{i=0}^k \frac{n_i}{n} \cdot (G - 1) \quad (1)$$

2) Ecualización adaptativa [10]

Contraste adaptativo limitado de la ecualización del histograma (CLAHE), a diferencia de la ecualización del histograma, es un método local que mejora partes pequeñas de las imágenes llamadas mosaicos. Luego, cada uno de estos bloques se ecualiza como se hace normalmente con el algoritmo de ecualización del histograma tradicional. El objetivo de utilizar esta técnica es resaltar los componentes principales del rostro

abordando el problema de la iluminación por zonas. En la figura 3.2 se muestra una comparación entre la ecualización del histograma tradicional y la ecualización adaptativa. Como se puede observar, al trabajar en ventanas (mosaicos) las zonas pequeñas no son influidas por las áreas de tonalidad mayor, lográndose un mayor contraste.

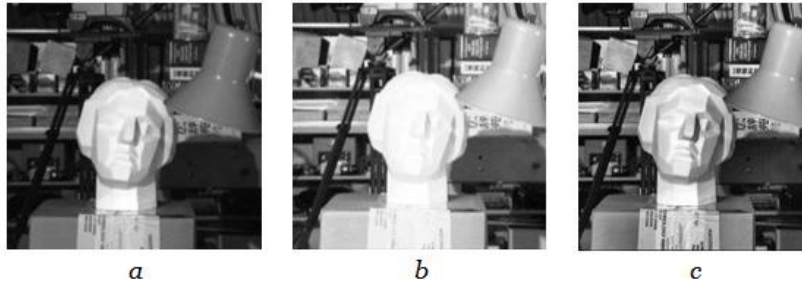


Figura 3.2 (a) Imagen en escala de grises, (b) ecualización del histograma tradicional, (c) ecualización adaptativa. [33]

3) Modelo de retina [10] [19]

El modelado de retina se emplea para lidiar con condiciones de iluminación desiguales, imitando el desempeño de la retina humana. Esta técnica de pre-procesamiento está compuesta de tres capas:

- Capa de foto receptores.
- Capa plexiforme externa (OPL).
- Capa plexiforme interna (IPL).

La capa de foto receptores tiene como objetivo imitar la propiedad de los bastones y conos de la retina humana, los cuales tienen la capacidad de resaltar los detalles por la noche o bajo señales brillantes en el día, sirviendo como filtros de adaptación a la luz. Para explotar y reemplazar esta habilidad, se usa una función no lineal basada en la ecuación de Naka-Rushton [34], como se muestra en la ecuación 2.

$$Y = \frac{X}{X + X_0} \quad (2)$$

Donde X representa los valores de entrada y X_0 es un factor de adaptación, que varía para aumentar la sensibilidad del método. Para modelar el procesamiento dentro de la capa OPL para la normalización de la iluminación, la imagen de entrada pasa a través de dos filtros de paso bajo gaussianos y un filtro de diferencias gaussianas (DoG). El uso de filtros de diferentes tamaños permite la eliminación de ruidos. Por último se pasa por la capa IPL para mejorar la apariencia global durante el mapeo de tonos, esta capa tiene dos funciones; una es proporcionar iluminación local a la capa OPL, la otra es realizar la mejora del contorno.

4) Retinex multi-escala [10] [16] [18] [19]

Modelo de luminosidad que logra la compresión simultánea del rango dinámico, la consistencia del color y la interpretación de la luminosidad con el objetivo de mejorar la fidelidad de las imágenes en color, a la observación humana.

Las imágenes con grandes discontinuidades de iluminación son susceptibles a los efectos de halo después de ser procesadas por el enfoque de retinex de escala única (SSR). Para superar este problema se creó una extensión del algoritmo SSR llamado retinex multi-escala (MR), que genera una combinación de imágenes producidas por SSR pero con diferentes anchos gaussianos. La reflectancia en el enfoque de MR se define por:

$$R(x, y) = \sum_{n=1}^N \omega_n \{ \log[I(x, y)] - \log[F(x, y) * I(x, y)] \} \quad (3)$$

Donde $\sum_{n=1}^N \omega_n$ representa la suma ponderada de los filtros gaussianos con diferentes escalas N y un factor de ponderación ω_n .

3.2 Algoritmo propuesto: foto receptores con retinex multi-escala (FRRM)

Se realizó una fusión entre la capa de foto receptores del algoritmo de retina y el algoritmo retinex multi-escala, esto para lograr la homogeneidad en las imágenes que proporciona el modelo de retina y resaltar los puntos característicos del rostro mediante el contraste que proporciona el algoritmo retinex.

El algoritmo fusiona la capa de foto receptores, utilizada en el algoritmo de retina, los foto receptores pueden ajustar su sensibilidad con respecto a la luminosidad de su vecindad. Esto se modela mediante la relación de *Michaelis-Menten* [35], la cual se normaliza para un rango de luminosidad de $[0, V_{max}]$ (ver ecuación 4). V_{max} Representa el valor máximo de píxeles en la imagen, 255 en el caso de imágenes estándar de 8 bits.

$$C(p) = \frac{R(p)}{R(p) + R_0(p)} (V_{max} + R_0(p)) \quad (4)$$

La salida $C(p)$ del foto receptor p depende del valor de $R(p)$ y de un parámetro de compresión $R_0(p)$ que está ligado linealmente a la iluminación local $L(p)$ (ver ecuación 5) del vecindario del foto receptor p . Esta iluminación local $L(p)$ se calcula aplicando un filtro de paso bajo espacial a la imagen de entrada.

$$R_0(p) = V_0 \cdot L(p) + V_{max}(1 - V_0) \quad (5)$$

La figura 3.3 muestra la aplicación de la capa de foto receptores a una imagen con mala y buena iluminación. Como resultado, se mejora la visibilidad del contraste en áreas oscuras (a), mientras las áreas con mayor iluminación resultan con menor afectación (b).



Figura 3.3 a) Imagen de entrada con mala iluminación, b) imagen con buena iluminación. La imagen de salida muestra el resultado de aplicar la capa de foto receptores a la imagen de entrada.

Hasta este punto el modelo de retina ayuda a iluminar la imagen de una forma homogénea; sin embargo, se necesita un mayor contraste en las zonas de interés del rostro (ojos, nariz y boca) para lo cual se agregaron características del algoritmo retinex multi-escala, el cual originalmente es utilizado en imágenes acuáticas para visualizar formas, que a simple vista no son perceptibles, logrando un mayor contraste.

El contraste se obtiene mediante el uso de envolventes gaussianas de diferentes anchos para asegurar una compresión del rango dinámico y un rendimiento en el ajuste tonal adecuado. La forma del algoritmo retinex multi-escala (MR) está descrito en la ecuación 3.

Para el funcionamiento del retinex multi-escala la imagen se divide en tres canales, cada canal es tratado por separado con el filtro gaussiano; al final del proceso, las tres imágenes se fusionan, obteniendo una imagen con un mejor rango dinámico y saturación. En la figura 3.4 se puede observar el resultado visual de aplicar a una imagen con mala iluminación los algoritmos de retina, retinex multi-escala y el algoritmo propuesto (FRRM).



Figura 3.4 Resultado visual de la aplicación de los algoritmos listados: a) Imagen de entrada, b) retina, c) retinex multi-escala, d) técnica propuesta FRRM.

3.3 Descripción del rostro

Se utiliza la librería Dlib [36] para la detección y descripción del rostro en términos de puntos faciales.

Dlib es una librería de software de código abierto multiplataforma escrita en el lenguaje de programación C ++. Su diseño está influenciado por el diseño por contrato y la ingeniería de software basada en componentes. Esto significa que es una colección de componentes de software independientes, cada uno acompañado de una extensa documentación y modos de depuración exhaustivos. Además, la librería está diseñada para ser útil tanto en proyectos de investigación como en proyectos comerciales del mundo real y se ha diseñado para que sea fácil de integrar en la aplicación C ++ de un usuario.

Su desarrollo comenzó en 2002, crecido para incluir una amplia variedad de herramientas. En 2009 Dlib se publicó en el *Journal of Machine Learning Research* [36]. A partir de 2016, contiene componentes de software para tratar redes, hilos, interfaces gráficas de usuario, estructuras de datos, álgebra lineal, aprendizaje automático, procesamiento de imágenes, minería de datos, análisis de XML y texto, optimización numérica, redes bayesianas y muchas otras tareas. En los últimos años, gran parte del desarrollo se ha centrado en la creación de un amplio conjunto de herramientas estadísticas de aprendizaje automático.

3.4 Algoritmos de clasificación

Esta etapa tiene como objetivo verificar que las imágenes capturadas durante el monitoreo en línea correspondan a uno de los aspirantes registrados. En la literatura se encontró que se emplean diferentes clasificadores para abordar el problema de la verificación de identidad, derivado de esto se realizaron pruebas con los siguientes 4 clasificadores:

- Máquinas de vector soporte (SVM) [12], con los siguientes parámetros:
 - Kernel: Polinomial
 - DEGREE: 1
 - GAMMA: 0.006
 - COEF: 0.1
 - NU: 1
 - P: 0.5
 - C: 3000
- Redes bayesianas [37]
- Perceptrón multicapa [18]
 - Capas: 100
- Random forest
 - Profundidad: 2000

Los valores de los parámetros de los clasificadores se obtuvieron después de realizar pruebas con variaciones de estos. En el caso de las máquinas de vector soporte las pruebas comprenden diferentes Kernel: Polinomial, sigmoïdal, radial y lineal. Así como variaciones en el rango de los valores DEGREE, GAMMA, COEF, UN, P y C, para encontrar el punto más alto de estabilidad del clasificador respecto al vector de características. En el clasificador de perceptrón multicapa, de igual manera, se realizaron diversas pruebas respecto a la cantidad de capas ocultas a utilizar, en tanto con random forest se utiliza una profundidad de 2000 para tener robustez ante las variaciones en número de clases.

La finalidad de evaluar diferentes clasificadores en la verificación de identidad es revisar el comportamiento del vector de características, debido a que, aunque el pre-procesamiento mejora la calidad de la imagen para lograr que la máscara se ajuste con mayor exactitud al rostro, la relevancia de las distancias faciales puede variar en los clasificadores.

En la figura 3.7 se muestra el objetivo del clasificador dentro del sistema para la verificación de identidad. En la etapa de entrenamiento se procesaran todas las imágenes de los aspirantes a presentar el examen que el monitor ingresa al sistema, una vez que se genera el conjunto de entrenamiento se ingresa al clasificador para generar el modelo del entrenamiento el cual crea un archivo .xml como se ilustra en la figura 3.5 en el apartado de entrenamiento. Después de crear el modelo de clasificación

se puede proceder al monitoreo, en el cual las imágenes se irán procesando una a una de acuerdo a la captura de imagen de cada aspirante, cuando se tiene el vector de característica de la imagen a analizar se introduce al modelo previamente guardado(.xml) para realizar la verificación de identidad.

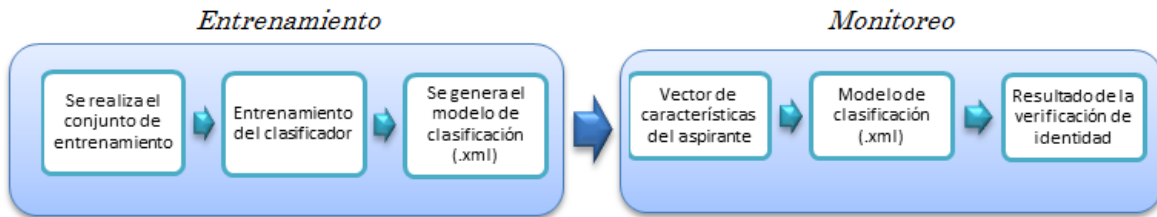


Figura 3.5 Objetivo del algoritmo de clasificación en el sistema.

3.5 Métricas

SSIM

Para evaluar la similitud de las imágenes en condiciones de iluminación homogénea y mala iluminación se utiliza la métrica de índice de similitud estructural (SSIM) [38] la cual proporciona información de la calidad en la similitud estructural, desde la perspectiva de la formación de imágenes. La ecuación se muestra a continuación:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (6)$$

Donde $x = \{x_i | i = 1, 2, \dots, N\}$ y $y = \{y_i | i = 1, 2, \dots, N\}$ son dos señales discretas no negativas que están alineadas entre sí, y donde μ_x , σ_x^2 y σ_{xy} son la media de x , la varianza de x y la covarianza de x y y , respectivamente. Además, μ_x y σ_x se pueden ver como estimaciones de la luminancia y el contraste de x , y σ_{xy} que miden la tendencia de x e y a variar entre sí, por lo tanto indica la similitud estructural.

Matriz de confusión

Para evaluar el modelo de clasificación se utiliza la matriz de confusión [39], que permite conocer la distribución del error a lo largo de las clases cuando se prueba un clasificador en un conjunto de datos que no intervienen en el entrenamiento. La estructura general de la matriz de confusión se muestra en la figura 3.8:

		Predicción	
		Positivos	Negativos
Observación	Positivos	Verdaderos Positivos (VP)	Falsos Negativos (FN)
	Negativos	Falsos Positivos (FP)	Verdaderos Negativos (VN)

Figura 3.6 Estructura general de la matriz de confusión [39].

Los valores que se encuentran a lo largo de la diagonal principal de la matriz, representan las clasificaciones correctas y los que están a lo largo de la diagonal secundaria representan los errores (la confusión) entre las clases, en donde:

- **VP** es la cantidad de positivos que fueron clasificados correctamente como positivos por el modelo.
- **VN** es la cantidad de negativos que fueron clasificados correctamente como negativos por el modelo.
- **FN** es la cantidad de positivos que fueron clasificados incorrectamente como negativos.
- **FP** es la cantidad de negativos que fueron clasificados incorrectamente como positivos.

Basado en la matriz de confusión se calculan las métricas de exactitud (ecuación 7), tasa de error (ecuación 8), sensibilidad (ecuación 9), especificidad (ecuación 10) y precisión (ecuación 11), las cuales se describen brevemente a continuación.

Exactitud

Porcentaje general de los datos clasificados correctamente. Se refiere a la dispersión del conjunto de valores, cuanto menor es la dispersión mayor es la exactitud. Se representa por la proporción entre el número de predicciones correctas (tanto positivas como negativas) y el total de predicciones.

$$Exactitud = \frac{VP + VN}{VP + VN + FN + FP} \quad (7)$$

Tasa de error

Porcentaje de datos clasificados incorrectamente.

$$Tasa\ de\ error = \frac{FP + FN}{VP + VN + FN + FP} \quad (8)$$

Sensibilidad

También se conoce como tasa de verdaderos positivos. Es el porcentaje de casos positivos que fueron correctamente identificadas por el algoritmo, indicando la capacidad del algoritmo para discriminar los casos positivos, de los negativos.

$$\text{Sensibilidad} = \frac{VP}{VP + FN} \quad (9)$$

Especificidad

Porcentaje de casos negativos que el algoritmo ha clasificado correctamente, nos indican la capacidad del algoritmo para discriminar los casos negativos, de los positivos.

$$\text{Especificidad} = \frac{VN}{FP + VN} \quad (10)$$

Precisión

Porcentaje de clasificación correcta cuando se predice positivo. Se refiere a lo cerca que está el resultado de una medición del valor verdadero. Se representa por la proporción entre los positivos reales predichos por el algoritmo y todos los casos positivos.

$$\text{Precisión} = \frac{VP}{VP + FP} \quad (11)$$

CAPÍTULO 4

Diseño del sistema

En este capítulo se describe el análisis y diseño del proceso necesario para el desarrollo del sistema de reconocimiento automático del rostro, para la verificación de identidad, para su evaluación en línea; este proceso incluye: el funcionamiento del sistema y las herramientas utilizadas (software y equipo) para el desarrollo del sistema.

4.1 Funcionamiento del sistema

La finalidad del sistema es verificar la identidad de los aspirantes que presentan examen de admisión en línea en ambientes no controlados. La arquitectura del sistema se muestra en la figura 4.1.

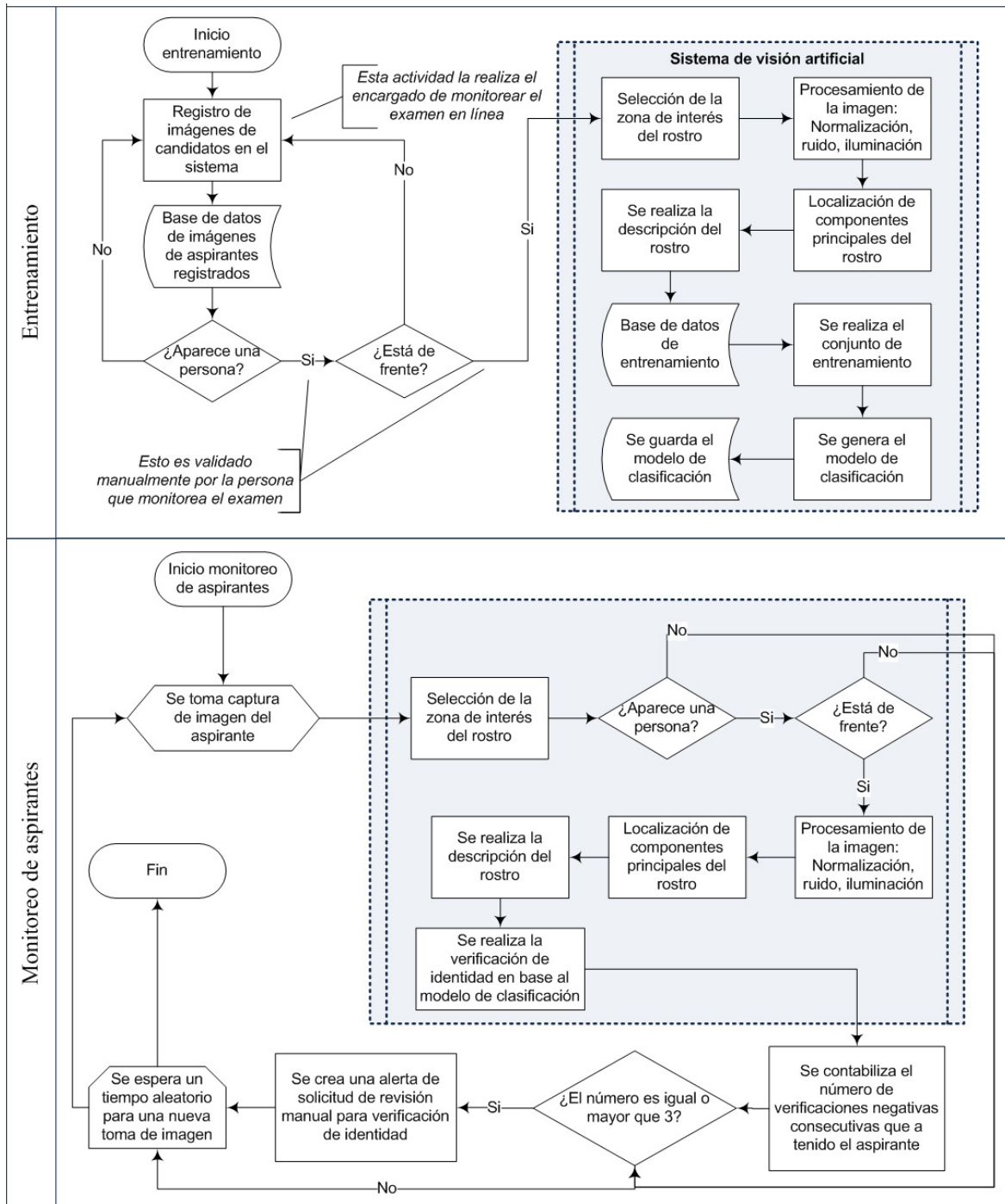


Figura 4.1 Diagrama del proceso del sistema.

A continuación se describe brevemente el proceso del sistema.

Fase de entrenamiento

- El encargado de monitorear el examen en línea, recibe las imágenes que los aspirantes envían mediante correo electrónico.
- Revisa que las imágenes correspondan a una cara de frente y se encarga de guardar la imagen en el área correspondiente y dar por bueno el registro del aspirante.
- El sistema de visión artificial toma el banco de imágenes de entrada e inicia el proceso; es decir,
 - El sistema buscar tener un rostro de frente.
 - Una vez que se localiza el rostro, se aplica el algoritmo de pre-procesamiento propuesto FRRM, para disminuir los efectos de iluminación propios del ambiente no controlado.
 - Se localizan los puntos de interés del rostro para calcular las distancias faciales que conforman el vector de características. Este proceso se repite para todos los registros de aspirantes que se tenga y se crea el banco de datos para la etapa de entrenamiento.
 - Posteriormente, se realiza el entrenamiento del clasificador para generar el modelo que es utilizado durante el monitoreo de los aspirantes.

Monitoreo de aspirantes

- Cuando se inicia el monitoreo de los aspirantes el sistema realiza capturas aleatorias, por medio de la cámara web, de cada aspirante registrado.
- La imagen del aspirante se introduce al sistema de visión artificial.
 - El sistema valida que la imagen capturada contenga a) al menos un rostro y que b) se encuentre de frente a la cámara, de no cumplir alguna de estas condiciones se realizará otra captura.
 - A la imagen que cumpla las condiciones anteriores, se le aplica el algoritmo de pre-procesamiento FRRM para disminuir los efectos de iluminación propios del ambiente no controlado.
 - Se localiza los puntos de interés del rostro para calcular las distancias faciales que conforman el vector de características de la imagen capturada.
 - Se realiza la verificación de identidad en base al vector de características obtenido y al modelo de clasificación generado en la fase de entrenamiento.

- El sistema contabiliza el número de verificaciones negativas consecutivas que tiene el aspirante, de ser igual a 3 se enviará una alerta al monitor del examen para que realice una verificación manual.

4.2 Descripción del rostro

Se utiliza la librería Dlib [36][40] en OpenCV [41] para la obtención de las partes relevantes de la cara, ojos, boca, nariz, etc.

Dlib detecta estructuras faciales importantes en la cara utilizando métodos de predicción de formas. La detección de puntos de referencia faciales es un proceso de dos pasos:

1. Localiza el rostro de una persona en la imagen.
2. Detectar las estructuras faciales clave del rostro.
 - a. Boca
 - b. Ceja derecha
 - c. Ceja izquierda
 - d. Ojo derecho
 - e. Ojo izquierdo
 - f. Nariz
 - g. Mandíbula

El detector de puntos de referencia faciales incluido en la librería Dlib es una implementación del artículo *One Millisecond Face Alignment with an Ensemble of Regression Trees* [40]. Este método utiliza un conjunto de entrenamiento de puntos de referencia faciales etiquetados en una imagen. Estas imágenes se etiquetan manualmente, especificando coordenadas (x, y) de las regiones que rodean cada estructura facial.

Dados estos datos de entrenamiento, un conjunto de árboles de regresión se entrena para estimar las posiciones de puntos faciales directamente desde las intensidades de píxeles (es decir, no se está realizando una "extracción de características"). Esto da como resultado un modelo entrenado con un conjunto de puntos preestablecidos propio de Dlib; además, de permitir entrenar otros modelos; en la figura 4.2 se muestra el modelo proporcionado por dicha librería. En este caso se utiliza el que tiene por *default*, con la localización de 68 puntos faciales.

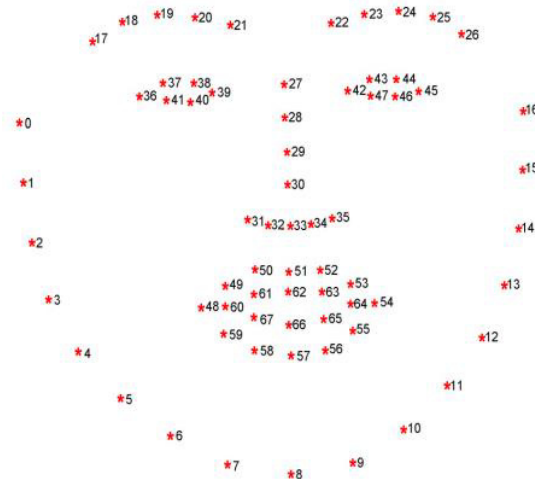


Figura 4.2 Puntos preestablecidos del rostro [42].

De los 68 puntos contenidos en la máscara, en este trabajo se utilizan 16 puntos para conformar 36 distancias faciales (ver figura 4.3) que son menos susceptibles a cambios debido a gestos faciales [43], estas distancias se normalizan para conformar el vector de características empleado en la etapa de verificación de identidad.

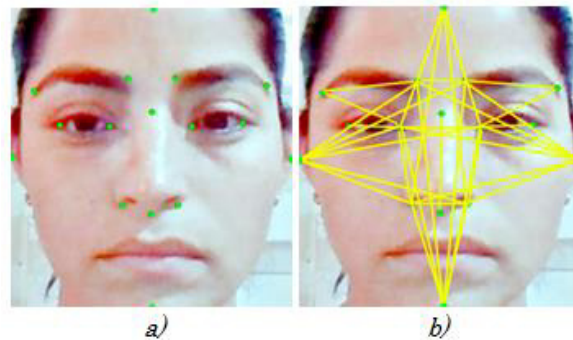


Figura 4.3 a) Puntos de interés del rostro, b) distancias faciales obtenidas a partir de los puntos de interés.

Las 36 distancias faciales que conforman el vector de características son:

1. Longitud ojo izquierdo
2. Longitud ojo derecho
3. Longitud fosas nasales
4. Largo nariz
5. Longitud ceja izquierda
6. Longitud ceja derecha
7. Distancia entre esquinas interna de cejas
8. Distancia entre esquinas internas de los ojos
9. Distancia entre esquina externa de ceja derecha a esquina interna de ojo derecho

10. Distancia entre esquina externa de ceja izquierda a esquina interna de ojo izquierdo
11. Distancia entre esquina interna de ojo derecho a fosa nasal derecha
12. Distancia entre esquina interna de ojo izquierdo a fosa nasal izquierda
13. Distancia entre esquina interna de ceja derecha a barbilla
14. Distancia entre esquina interna de ceja izquierda a barbilla
15. Distancia entre esquina interna de ojo derecho a barbilla
16. Distancia entre esquina interna de ojo izquierdo a barbilla
17. Distancia entre fosa nasal derecha a barbilla
18. Distancia entre fosa nasal izquierda a barbilla
19. Distancia entre fosa nasal derecha al oído derecho
20. Distancia entre fosa nasal izquierda al oído derecho
21. Distancia entre esquina interna de ojo derecho al oído derecho
22. Distancia entre esquina interna de ojo izquierdo al oído derecho
23. Distancia entre esquina interna de ceja derecha al oído derecho
24. Distancia entre esquina interna de ceja izquierda al oído derecho
25. Distancia entre fosa nasal derecha al oído izquierdo
26. Distancia entre fosa nasal izquierda al oído izquierdo
27. Distancia entre esquina interna de ojo derecho al oído izquierdo
28. Distancia entre esquina interna de ojo izquierdo al oído izquierdo
29. Distancia entre esquina interna de ceja derecha al oído izquierdo
30. Distancia entre esquina interna de ceja izquierda al oído izquierdo
31. Distancia entre fosa nasal derecha a la parte superior del rostro
32. Distancia entre fosa nasal izquierda a la parte superior del rostro
33. Distancia entre esquina interna de ojo derecho a la parte superior del rostro
34. Distancia entre esquina interna de ojo izquierdo a la parte superior del rostro
35. Distancia entre esquina interna de ceja derecha a la parte superior del rostro
36. Distancia entre esquina interna de ceja izquierda a la parte superior del rostro

4.3 Interface del sistema

La interfaz del sistema se estableció de forma que sea fácil ingresar las imágenes para el entrenamiento, y la ubicación desde donde se lleva a cabo el monitoreo de cada aspirante. Se compone de dos secciones: entrenamiento y monitoreo de aspirantes. La primera sección (figura 4.4) se conforma de un apartado en donde se puede introducir el nombre del aspirante o bien algún número de control con el que se desee identificar al mismo, un botón de *Agregar* para insertar el registro; así como, una ventana en donde se muestra una visualización de la imagen de entrenamiento seleccionada, en la parte inferior se muestra un listado de los aspirantes registrados y un botón para iniciar el entrenamiento, el cual se activará cuando se hayan ingresado un mínimo de 10 aspirantes.

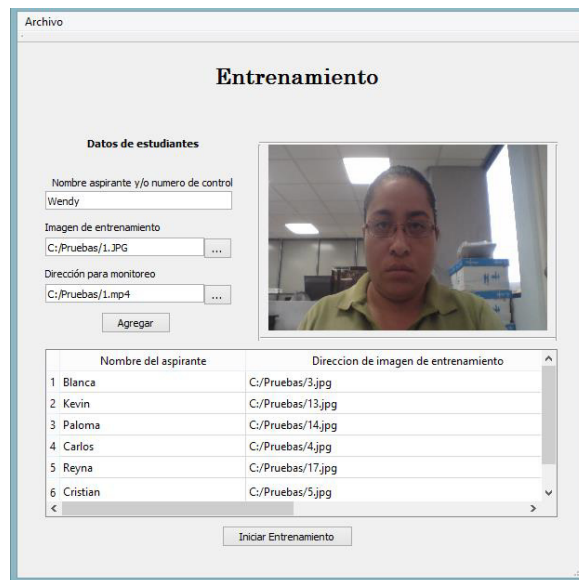


Figura 4.4 Interfaz de entrenamiento del sistema.

La segunda sección (figura 4.5) consta de un botón para iniciar el monitoreo y otro para detenerlo; así como, un apartado en donde se enlistan las alertas del sistema cuando la verificación de identidad, de un determinado aspirante, sea negativa en reiteradas ocasiones. Las alertas se notifican en pantalla cuando existen tres o más detecciones negativas (no se encuentra un rostro, no se encuentra en un ángulo visible frente a la cámara, el rostro no corresponde al aspirante que debe estar frente a la cámara) del sistema, esto es para evitar la saturación en pantalla de alertas positivas, dejando así sólo las alertas de interés, que son la verificación incorrecta de un aspirante.

También se cuenta con un cuadro de vista previa en donde el monitor tiene la opción de seleccionar al aspirante que desea realizar una inspección manual. Esta sección no está ligada a la sección de alertas, debido a que si se ejecuta en un mismo proceso los dos dependerían de que uno u otro termine para poder continuar con el proceso, haciendo que existan momentos en donde se pierda la verificación. Por lo tanto, los procesos se ejecutan en dos hilos diferentes para que, la parte de monitoreo manual sea independiente al proceso de verificación de identidad en tiempo real.



Figura 4.5 Interfaz de monitoreo del sistema.

4.4 Herramientas utilizadas (software y equipo)

Para el desarrollo del sistema se utilizaron las siguientes herramientas (ver anexo A):

- C++ en Visual Studio 2015, como entorno de desarrollo [44].
- OpenCV 3.4.0, para el procesamiento de imágenes [41].
- Librería Dlib C++, para la localización del rostro y puntos de interés [42].
- Qt 4.8.1, para la creación de la interfaz gráfica [45]

CAPÍTULO 5

Pruebas y Resultados

Este capítulo describe las pruebas realizadas para evaluar los módulos de localización del rostro, pre-procesamiento de imágenes y la verificación de identidad. En el pre-procesamiento de imágenes se evalúan y comparan los cinco algoritmos de pre-procesamiento explicados en las secciones 3.1 y 3.2, que son: ecualización del histograma (EH), ecualización adaptativa (CLAHE), modelo de retina, retinex multi-escala y el algoritmo propuesto (FRRM). Así como, cuatro combinaciones entre ellos que son: EH + retina, EH + retinex multi-escala, CLAHE + retina y CLAHE + retinex multi-escala. La métrica utilizada para evaluar de manera cuantitativa estas técnicas es la SSIM.

Para la verificación de identidad se comparan cuatro clasificadores que son: máquinas de vector soporte (SVM), redes bayesianas, perceptrón multicapa y random forest. Los resultados son evaluados mediante las métricas de: exactitud, precisión, especificidad, sensibilidad y tasa de error.

5.1 Banco de imágenes

Para este proyecto se cuenta con tres bancos de imágenes: 1) FEI Face Database para pruebas en el pre-procesamiento de las imágenes y verificación de identidad, 2) Faces94 para pruebas con verificación de identidad y 3) un banco de imágenes propio para pruebas de verificación de identidad y del funcionamiento general del sistema. A continuación se describen los diferentes bancos de imágenes.

1) FEI Face Database [46]

El banco de imágenes *FEI Face Database* contiene un conjunto de imágenes faciales tomadas entre junio de 2005 y marzo de 2006 en el laboratorio de inteligencia artificial de la FEI en São Bernardo do Campo, São Paulo, Brasil. Todas las caras están representadas principalmente por los estudiantes y el personal, en la figura 5.1 se muestra un ejemplo de la variación de imágenes por individuo.



Figura 5.1 Variaciones de imágenes por individuo del banco de imágenes FEI [46]

Descripción del banco de imágenes:

- 14 imágenes de 200 personas, en total 2800 imágenes.
- Tamaño de 640 x 480 píxeles.
- Edades entre 19 y 40 años.
- 100 mujeres y 100 hombres.
- Fondo blanco.
- Cambios en la posición de la cara.
- Con variación de iluminación en cada persona.
- Cambios en la expresión facial.
- Formato RGB.
- Oclusión parcial (barba y lentes).

2) Faces94 [47]

Banco de imágenes que aborda la variación de la expresión facial, los sujetos se sientan a una distancia fija de la cámara y hablan mientras se toma una secuencia de imágenes. En la figura 5.2 se muestra un ejemplo de las imágenes contenidas en el banco de imágenes.



Figura 5.2 Muestra de imágenes contenidas en el banco de imágenes Faces94. [47]

Descripción del banco de imágenes:

- 20 imágenes de 152 individuos, en total 3040 imágenes.
- Resolución de 180 x 200 píxeles, formato vertical.
- 20 mujeres, 113 hombres.
- 132 personas jóvenes y 20 personas mayores.
- Fondo verde claro.
- Cambios menores en la posición de la cara.
- Sin variación de iluminación en cada persona.
- Cambios considerables en la expresión facial.
- Formato RGB.
- Oclusión parcial (con cabello, barba y lentes).

3) Banco de imágenes propio

Como ya se mencionó, la adquisición de imágenes se integra de 2 etapas: a) adquisición de imágenes para el entrenamiento y b) adquisición de imágenes para el monitoreo.

- Adquisición de imágenes para el entrenamiento: los aspirantes a presentar examen deben proporcionar una fotografía de ellos mostrándose de frente a la cámara, esta debe ser la misma que utilizarán durante la aplicación del examen. Dichas fotografías son revisadas y validadas por el monitor del examen para que sean cargadas al sistema en el apartado de *Entrenamiento*.
- Adquisición de imágenes para el monitoreo: las capturas de imágenes durante el monitoreo se realiza por la cámara web de los aspirantes. Para las pruebas en el sistema las imágenes son capturadas de los videos del banco de imágenes propio.

Para pruebas de verificación de identidad y del funcionamiento general del sistema se creó un banco de imágenes propio, en el mes de agosto del 2018, en ambientes naturales; consta de una imagen tomada mediante cámara web de personas de frente (para la fase de entrenamiento) y un vídeo de 5 minutos de duración, grabado con la misma cámara que se tomó la imagen inicial (para la etapa de monitoreo), en la figura 5.3 se muestran las capturas de las imágenes contenidas en el banco de imágenes.



Figura 5.3 Muestra de las imágenes contenidas en el banco de imágenes propio.

Descripción del banco de imágenes:

- 20 individuos.
- 1 imagen frontal para el registro inicial, por individuo.
- 1 video con duración de 5 minutos por individuo.
- Iluminación y pose variable durante el vídeo.
- 13 mujeres y 7 hombres.
- Resolución variable (301 x 157; 320 x 240; 640 x 360; 640 x 480; 1280 x 720).
- Formato RGB.
- Oclusión parcial (sombrero, audífonos y lentes).
- Escala e intensidad luminosa diferente.
- Fondos estáticos y dinámicos reales.

5.2 Experimentación

Se realizaron tres diferentes pruebas las cuales tienen por objetivo:

Localización del rostro

La finalidad de esta prueba es examinar el módulo de detección y localización del rostro; es analizar qué tan robusto es la librería Dlib ante imágenes que contengan “objetos” con características semejantes a las de un rostro de una persona. Además, examinar la detección del rostro a partir de diferentes ángulos del rostro frente a la cámara.

Pre-procesamiento de imágenes

Esta prueba tiene como objetivo evaluar de manera cuantitativa, mediante la métrica SSIM, el rendimiento de las distintas técnicas de procesamiento implementadas y la propuesta en este trabajo, denominada FRRM, ante una imagen con buena iluminación y una imagen con mala iluminación.

Verificación de identidad

El objetivo de esta prueba es realizar la verificación de la identidad de los aspirantes registrados, para ello se evalúa el funcionamiento de 4 algoritmos de reconocimiento de patrones. Se analizan los resultados mediante las métricas de exactitud, tasa de error, precisión y similitud.

5.3 Análisis de resultados

5.3.1 Localización



Las pruebas de localización del rostro se dividen en dos apartados.

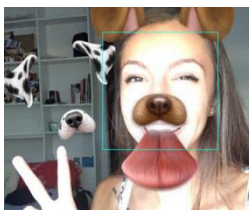
A) Prueba 1

Para la prueba 1 de localización del rostro se utilizaron imágenes de animales, robots, payasos, caricaturas, peluches así como una fotografía de una mujer con adornos de perro.

En la tabla 5.1 se listan las imágenes de prueba, en las cuales se busca el rostro de una persona, cuando se logra la detección, la cara se enmarca en un cuadro verde.

Tabla 5.1 Resultados de detección del rostro.

Imagen	Análisis del resultado
	En la imagen se visualiza un perro frente a la cámara de una computadora, como se puede observar la librería Dlib no detecta un rostro.
	Esta imagen al igual que la anterior es un animal sin embargo se le agregaron lentes, a pesar de esta característica la librería no detecta un rostro.



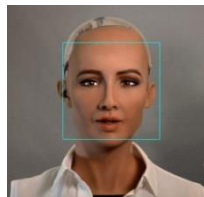
La imagen presenta un filtro de *Snapchat*; sin embargo, las características del rostro como son ojos y boca prevalecen, lo que hace que se logre la localización de la cara.



En la imagen se puede observar la presencia de una persona y un perro; no obstante, la librería detecta de manera correcta sólo el rostro de la persona.



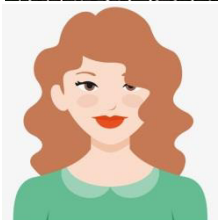
En imágenes que contienen rostros de robots no se realiza la detección del rostro, debido a que las características faciales presentes no son similares a una persona.



Al contrario de la imagen anterior, en esta imagen se muestra el robot *Sophia*, cuyos rasgos faciales imitan a las de un ser humano, por lo que la librería detecta el rostro.



En la imagen se muestra la cara de un payaso, la cual visualmente tiene los componentes principales de un rostro, pero como son alterados en tamaño debido al maquillaje, la librería no detecta el rostro.



En la imagen de una caricatura de una mujer tampoco se logra la detección del rostro.



Al igual que la imagen anterior, se presenta una fotografía que contiene un peluche cuyo rostro no cumple con los rasgos faciales de un ser humano, por lo que, no se detecta rostro.

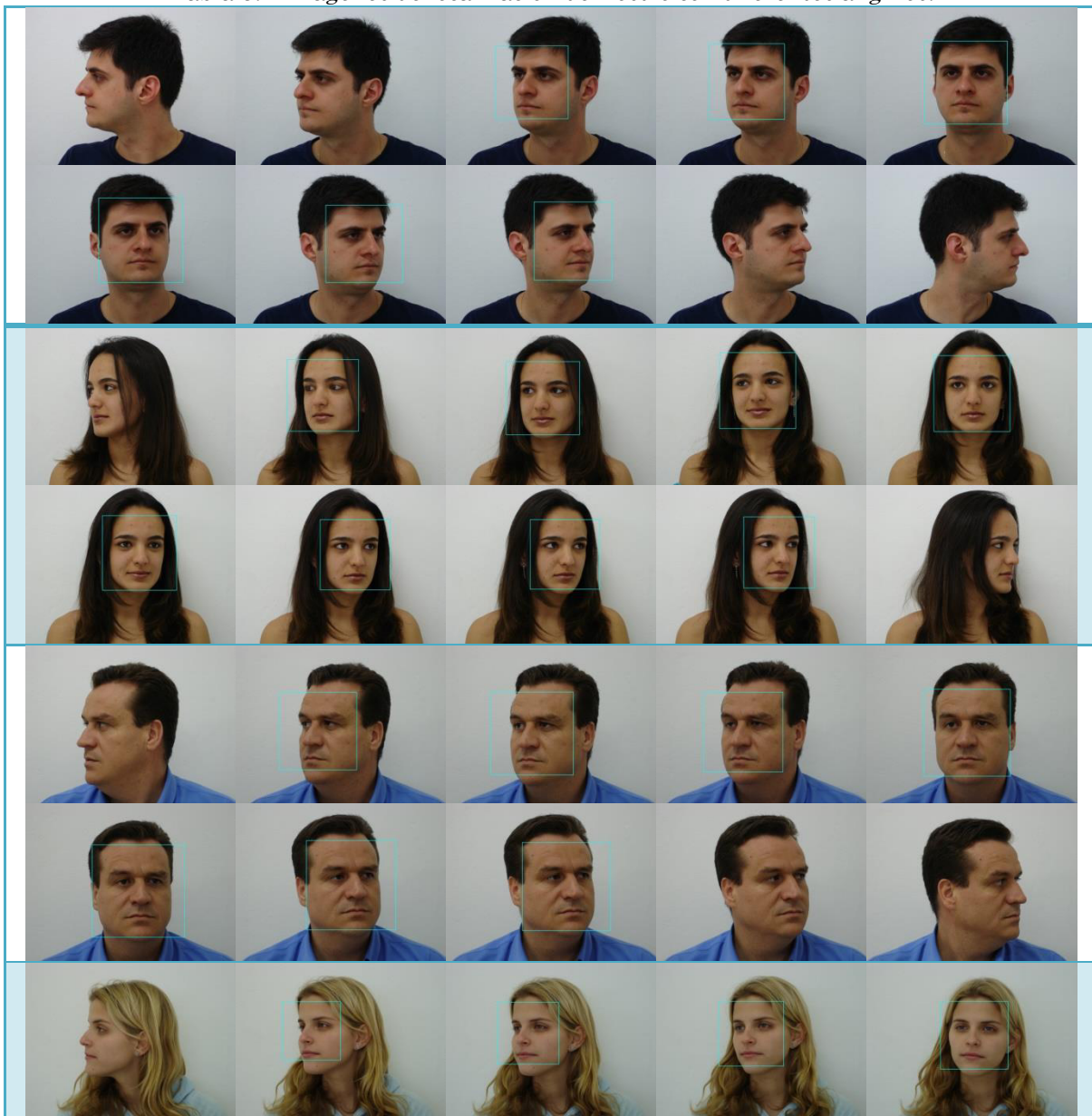
Comentarios

Con la evaluación realizada, se puede concluir, que la librería Dlib [42], tiene un buen desempeño al detectar caras que cumplen con las relaciones existentes entre los rasgos faciales del ser humano.

B) Prueba 2

Respecto a la localización del rostro, se utilizaron fotografías de personas, con 10 diferentes posiciones del rostro frente a la cámara web. La finalidad fue analizar la robustez de la librería Dlib [42] en la detección del rostro ante diferentes posiciones (ángulos) del rostro frente a una cámara web. En la tabla 5.2 se observan ejemplos de las 10 rotaciones (en un eje) del rostro que van desde el 0 hasta 180 grados, variando 20 grados, este se enmarca en un cuadro verde cuando se logra la detección.

Tabla 5.2 Imágenes de localización del rostro con diferentes ángulos.





Comentarios

Con la evaluación realizada, se puede concluir, que la librería Dlib [42], tiene un buen desempeño al realiza la detección de la cara, sólo cuando puede localizar todos los componentes faciales, que son: ojos, nariz y boca.

5.3.2 Pre-procesamiento

Para las pruebas se utiliza el banco de imágenes FEI Face Database, considerando por cada individuo una imagen con iluminación controlada y una imagen con variación en la iluminación, haciendo un total de 400 imágenes (200 personas diferentes). Como ya se comentó, el cambio en la intensidad luminosa es grande. La figura 5.4 muestra un subconjunto de 5 imágenes, las cuáles servirán para mostrar de manera visual los resultados de la aplicación (sobre ellas) de cada técnica.



Figura 5.4 (a) Imágenes con iluminación controlada, (b) imágenes con iluminación variable.

A continuación, se muestran las pruebas realizadas con las diferentes técnicas de pre-procesamientos, en imágenes con el rostro ya localizado. Las pruebas se dividen en 3 apartados.

A) Prueba 1

Esta prueba tiene como objetivo evaluar de manera separada a cada técnica de pre-procesamiento (1)ecualización del histograma, 2)ecualización adaptativa, 3)modelo de retina, 4)retinex multi-escala y 5)algoritmo propuesto). Así como la combinación directa de las técnicas retina y retinex multi-escala, ambas base del método propuesto. Los resultados obtenidos se analizan en términos de la métrica SSIM, en una gráfica comparativa.

1) Ecuación del histograma

En la figura 5.5 se muestra de manera visual el resultado obtenido. Se puede observar que funciona relativamente bien ante imágenes con diferente intensidad luminosa. Si bien se tiene una mejora en el contraste de las imágenes, los rasgos de interés no están definidos.



Figura 5.5 Imágenes resultantes de aplicar el filtro de ecuación del histograma a imágenes con iluminación controlada e iluminación variable.

2) Ecuación adaptativa (CLAHE)

En la figura 5.6 se observa que en las imágenes con iluminación controlada los rasgos tienen mayor definición; sin embargo, en imágenes oscuras la imagen no logra mejorar la calidad de la imagen para localizar los puntos de interés.



Figura 5.6 Imágenes resultantes de aplicar el filtro de ecuación adaptativa del histograma a imágenes con iluminación controlada e iluminación variable.

3) Modelo de retina

Con esta técnica, se observa que las imágenes resultantes tienen mayor nitidez en cuanto al rostro en general; sin embargo, la tonalidad es homogénea haciendo que los rasgos distintivos (ojos, nariz, boca) no estén completamente definidos.

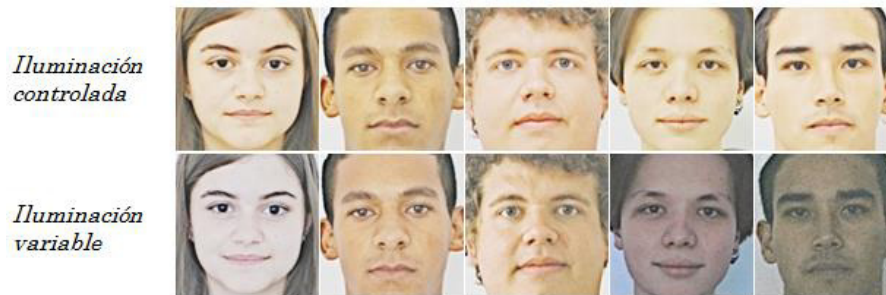


Figura 5.7 Imágenes resultantes de aplicar el filtro de modelo de retina a imágenes con iluminación controlada e iluminación variable.

4) Retinex multi-escala

En la figura 5.8 se muestra la comparativa de las imágenes (con iluminación controlada e iluminación variable) resultantes de haberles aplicado, la técnica de procesamiento retinex multi-escala. Visualmente, esta técnica logra un obtener un mayor contraste; sin embargo, en la zona de los ojos el contraste es mayor, haciendo que se pierda el detalle de éstos.



Figura 5.8 Imágenes resultantes de aplicar el filtro de retinex multi-escala a imágenes con iluminación controlada e iluminación variable.

5) Algoritmo propuesto (FRRM)

Se propuso el algoritmo FRRM debido a que se detectó que los métodos anteriores no obtienen un realce importante de las zonas de interés del rostro (ojos, nariz, boca). En este algoritmo se fusionan las características que proporciona la homogeneidad en el rostro por parte del modelo de retina (específicamente la capa de foto receptores) y el contraste que proporciona retinex multi-escala en las zonas deseada. Todo ello para lograr mayor definición visual como se muestra en la figura 5.9.

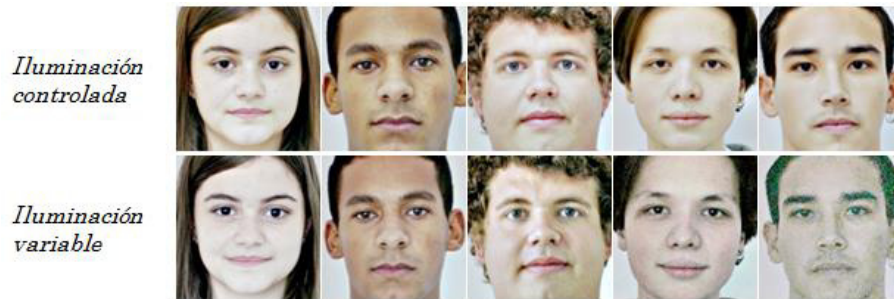


Figura 5.9 Imágenes, con iluminación controlada e iluminación variable, resultantes de haberles aplicado la técnica propuesta FRRM.

6) Combinación retina + retinex multi-escala

Se realiza una prueba con la combinación de las técnicas retina + retinex multi-escala para mostrar las diferencias respecto al algoritmo propuesto FRRM, los resultados se muestra en la figura 5.10. Como se puede ver, el realizar la combinación directa de ambas técnicas proporciona resultados visuales no tan buenos. En todos los casos, las imágenes son claras y no muestran un contraste adecuado.

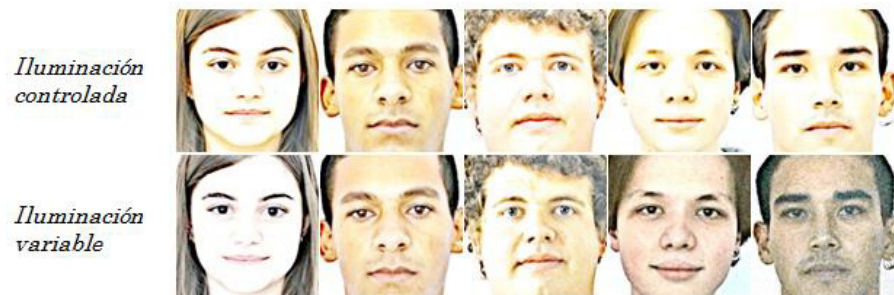


Figura 5.10 Imágenes con retina + reinex aplicado a imágenes con iluminación controlada e iluminación variable.

Comparativa de los métodos aplicados respecto a la métrica SSIM

Para evaluar los resultados de las imágenes de manera cuantitativa se calculó la métrica SSIM, comparando las imágenes de iluminación controlada (considerando que son éstas las que mejor contraste tienen) con respecto a las imágenes resultantes de las distintas pruebas. Como se muestra en la figura 5.11, el algoritmo de retina tiene el mejor porcentaje, en comparación a los otros algoritmos; sin embargo, a pesar de su semejanza de manera numérica, el problema a resolver sigue siendo la nitidez de las principales zonas del rostro para localizar de manera precisa los puntos característicos del rostro.

Como se puede apreciar en la gráfica (figura 5.11), el algoritmo FRRM queda en cuarto lugar, atrás de las técnicas de retina, la combinación de retina + retinex multi-escala y de la ecualización del histograma. Resultados que concuerdan con lo encontrado en el estado del arte.

Se esperaba lograr un mejor rendimiento con nuestro algoritmo; sin embargo, analizando los resultados de la métrica SSIM, los valores indican que hay una mayor diferencia de la imagen resultante con respecto a mejor iluminación, pero eso no significa que esta imagen, tenga las mejores condiciones para el siguiente proceso.

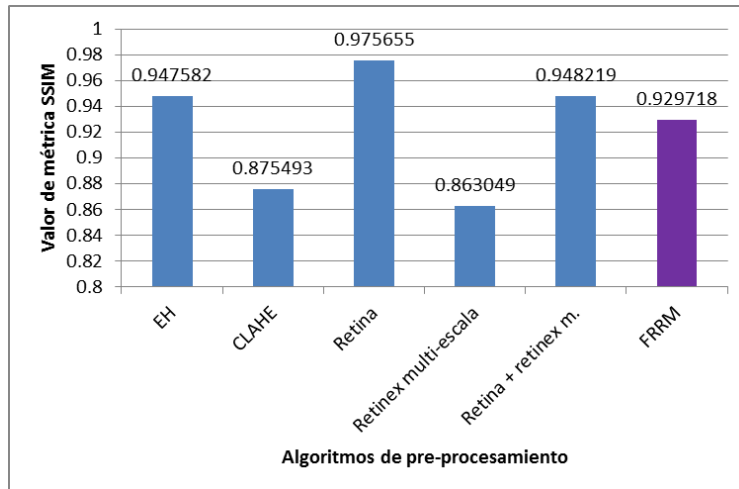


Figura 5.11 Gráfica comparativa de algoritmos de pre-procesamiento respecto a la métrica SSIM.

B) Pruebas 2

Si bien los resultados en la prueba 1 son buenos, de acuerdo a la literatura antes del pre-procesamiento, generalmente, se realiza una normalización de las tonalidades de grises para reducir el problema de la iluminación, por esta razón, se realizaron las siguientes pruebas que toma como base la ecualización del histograma y posteriormente se aplican los algoritmos de retina y retinex multi-escala. A continuación se muestran los resultados obtenidos, seguidos de una gráfica comparativa respecto a la métrica SSIM.

Ecualización del histograma + retina

En la figura 5.12 se muestra la comparativa visual de las imágenes (con iluminación controlada e iluminación variable) resultantes de haber procesado las imágenes con la ecualización del histograma y posteriormente aplicarles el algoritmo de retina. Se observa visualmente que las imágenes obtienen un tono homogéneo sin mayor realce, no logrando contrastar las zonas correspondientes a los ojos, nariz y boca.



Figura 5.12 Imágenes con iluminación controlada e iluminación variable, resultantes de aplicarles una ecualización del histograma y posteriormente el algoritmo de retina.

Ecualización del histograma + retinex multi-escala

En la figura 5.13 se muestra la comparativa visual de las imágenes (con iluminación controlada e iluminación variable) resultantes de haber procesado las imágenes con la ecualización del histograma y posteriormente aplicarles el algoritmo retinex multi-escala. Se observa visualmente que las imágenes tienen un mayor contraste; sin embargo, en las personas de tez morena el contraste difumina los rasgos de interés del rostro (ojos, nariz y boca).



Figura 5.13 Imágenes con iluminación controlada e iluminación variable, resultantes de aplicarles una ecualización del histograma y posteriormente el algoritmo de retinex multi-escala.

Comparativa de los métodos aplicados respecto a la métrica SSIM

Como se muestra en la figura 5.14, el algoritmo de retina sigue logrando el mejor porcentaje, en comparación al otro algoritmo. La diferencia numérica es significativa; sin embargo, la combinación de ecualización del histograma con retina da como resultado (combinación con el mejor rendimiento), imágenes en donde rasgos importantes para la descripción de la cara como los ojos, no están visualmente definidos.

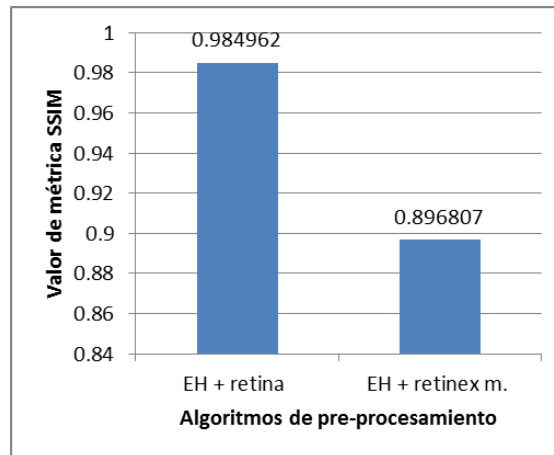


Figura 5.14 Gráfica comparativa de los métodos ecualización de histograma con retina y ecualización de histograma con retinex multi-escala, respecto a la métrica SSIM.

C) Pruebas 3

Derivado de la prueba 2 se realizó otra evaluación, pero en esta prueba se toma como base la ecualización adaptativa del histograma y posteriormente se aplican los algoritmos de retina y retinex. Seguido de una gráfica comparativa respecto a la métrica SSIM.

Ecualización adaptativa + retina

En la figura 5.15 se muestra la comparativa visual de las imágenes con iluminación controlada e iluminación variable con un procesamiento de retina, después de haber procesado las imágenes con la ecualización adaptativa del histograma. Se observa que las imágenes presentan mayores contrastes con realce en las zonas de interés; no obstante, en las imágenes de iluminación variable, más oscuras, no logra obtener la claridad o contraste suficiente en las zonas de interés.



Figura 5.15 Imágenes con iluminación controlada e iluminación variable, resultantes de aplicarles una ecualización adaptativa y posteriormente el algoritmo de retina.

Ecualización adaptativa + retinex multi-escala

En la figura 5.16 se muestra la comparativa de las imágenes con iluminación controlada e iluminación variable con un procesamiento de retinex multi-escala,

después de haber procesado las imágenes con la ecualización adaptativa del histograma. Se observa que las imágenes tienen mayor contraste en las zonas de interés; sin embargo en algunos casos aún se presentan en las imágenes sombras en los ojos y destellos de iluminación que hacen que se pierda el detalle de las zonas de interés.



Figura 5.16 Imágenes con iluminación controlada e iluminación variable, resultantes de aplicarles una ecualización adaptativa y posteriormente el algoritmo de retinex multi-escala.

Comparativa de los métodos aplicados respecto a la métrica SSIM

Como se muestra en la figura 5.17, de acuerdo con la métrica SSIM el algoritmo de retina sigue teniendo el mejor porcentaje, esto indica que las imágenes con iluminación variable tienen mayor semejanza con las imágenes de iluminación controlada en comparación al otro algoritmo.

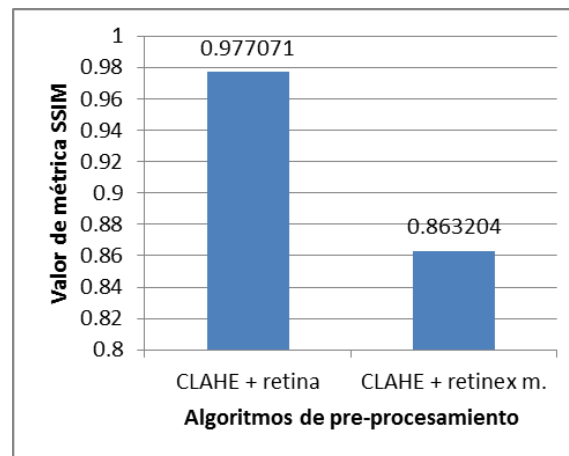


Figura 5.17 Gráfica comparativa de los métodos ecualización adaptativa con retina y ecualización adaptativa con retinex multi-escala, respecto a la métrica SSIM.

Comentarios finales de las pruebas de pre-procesamiento

Como se muestra en la figura 5.18, de manera cuantitativa la combinación del algoritmo de ecualización del histograma con el de retina tiene el mejor porcentaje, en comparación a los otros algoritmos (barra verde), es decir, esta combinación logra una imagen con características similares a las que tiene la imagen con una buena iluminación. Como se puede apreciar en la gráfica, el algoritmo FRRM queda en sexto

lugar (barra en morado) atrás de la combinación de ecualización del histograma con retina, ecualización adaptativa con retina, retina, la combinación de retina + retinex multi-escala y de la ecualización del histograma.

Sin embargo, el problema a resolver sigue siendo la nitidez de las principales zonas del rostro para localizar de manera precisa los puntos característicos del rostro. A su favor, visualmente el algoritmo propuesto FRRM obtiene los mejores resultados para detectar los puntos característicos.

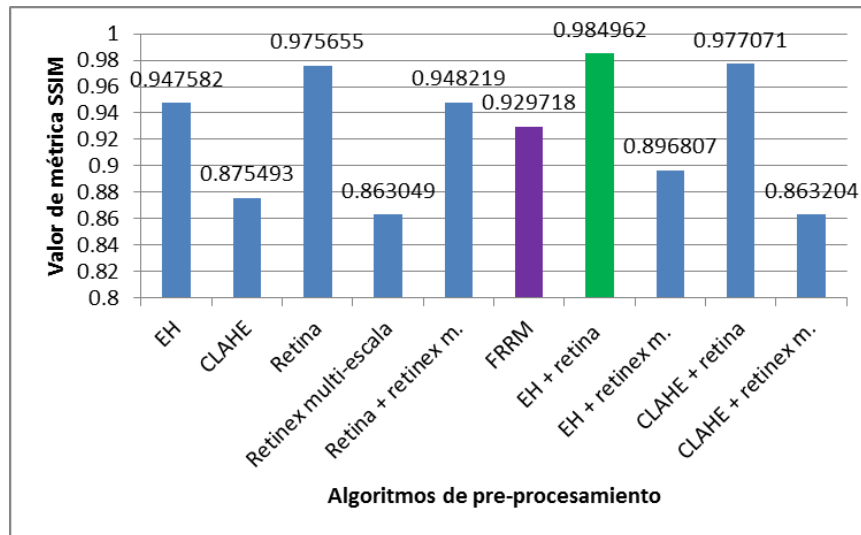


Figura 5.18 Grafica comparativa de los algoritmos de pre-procesamiento respecto a la métrica SSIM.

5.3.3 Verificación de identidad

Como ya se mencionó, el objetivo de esta etapa es verificar el rendimiento del modelo de clasificación obtenido en la etapa de pre-procesamiento. El conjunto de imágenes utilizado es diferente al correspondiente a la etapa de entrenamiento. Para las pruebas se utilizaron los tres bancos de imágenes *FEI Face Database*, *Faces94* y el banco de imágenes propio; considerando una imagen por aspirante para entrenamiento y el resto para la verificación de identidad.

A continuación, se muestran las pruebas divididas por cada banco de imágenes, en las cuales son comparados con los diferentes algoritmos de pre-procesamientos y con la combinación retina + retinex multi-escala para revisar la diferencia entre esta combinación y el algoritmo FRRM. Los resultados son evaluados con las métricas de exactitud, precisión, especificidad, sensibilidad y tasa de error.

FEI Face Database [46]

Para las pruebas se utilizaron 200 imágenes de entrenamiento correspondientes a una por aspirante (clase), así como 5 imágenes por clase haciendo un total de 1000 imágenes para la predicción, en la figura 5.19 se visualiza, una muestra, de las imágenes de entrenamiento y evaluación pertenecientes a este banco de imágenes.



Figura 5.19 Muestra de la imagen de entrenamiento y las correspondientes a la etapa de evaluación del banco de imágenes FEI Face Database.

En la tabla 5.3 se muestra la matriz de confusión de los resultados obtenidos con el algoritmo de pre-procesamiento propuesto (FRRM) y los diferentes clasificadores. En ella se observa que el porcentaje más alto de exactitud es 81.12% con el clasificador SVM.

Tabla 5.3 Resultados de clasificación del banco de imágenes FEI Face Database en combinación con el algoritmo de pre-procesamiento propuesto FRRM.

	Exactitud	Tasa de error	Sensibilidad	Especificidad	Precisión
SVM	81.12	18.77	81.22	99.88	83.78
Redes bayesianas	75.43	24.55	75.45	99.84	79.87
MLP	51.37	48.35	51.65	99.53	55.81
Random forest	76.85	22.98	77.03	99.85	79.85

La tabla 5.4 muestra los resultados de la métrica exactitud obtenida por todos los algoritmos de pre-procesamiento. En ella se observa, con letras verdes, los porcentajes más altos alcanzados por los algoritmos. El valor más alto es de 81.22% con el algoritmo EH + retina y el clasificador SVM, teniendo una diferencia de 0.1 respecto al algoritmo propuesto FRRM. Así también la diferencia entre el algoritmo FRRM y la combinación de los algoritmos retina+ retinex multi-escala es de 3.25 %.

Tabla 5.4 Resultados de exactitud del banco de imágenes FEI Face Database respecto a los algoritmos de pre-procesamiento.

	FRRM	Retina	Retinex m.	EH	CLAHE	EH + retina	EH + retinex m.	CLAHE + retina	CLAHE + retinex m.	Retina+ retinex m.
SVM	81.12	78.38	78.07	78.17	74.31	81.22	70.15	78.27	74.01	77.87
Redes bayesianas	75.43	71.07	68.32	70.76	66.9	73.3	65.69	67.41	64.77	71.37
MLP	51.37	49.75	44.57	50.25	46.5	50.96	43.65	49.24	41.22	46.7
Random forest	76.85	76.75	72.28	72.49	72.49	75.33	69.34	74.42	71.98	74.42

En la figura 5.20 se muestra la gráfica comparativa de los valores más altos de exactitud listados en la tabla 5.4, de la cual se puede concluir que el algoritmo de clasificación SVM tiene el mejor rendimiento.

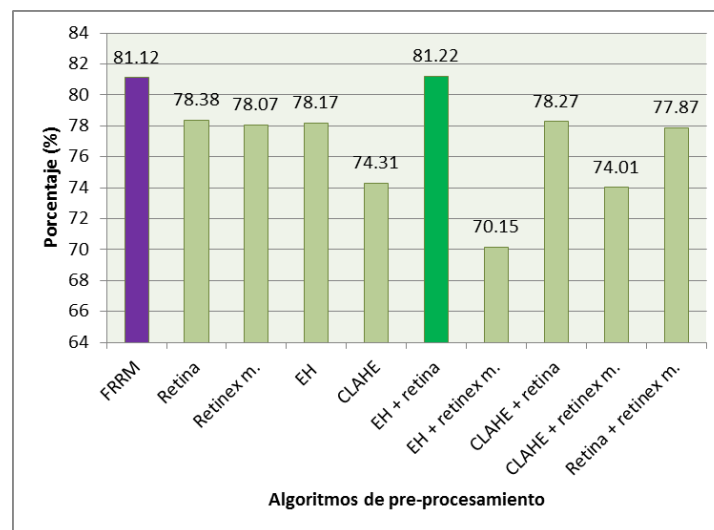


Figura 5.20 Comparativa de exactitud de los algoritmos de pre-procesamiento respecto al banco de imágenes FEI Face Database.

Faces94 [47]

Para las pruebas se utilizaron 152 imágenes de entrenamiento correspondientes a una por persona (clase), así como 20 imágenes por clase para la predicción, haciendo un total de 3040. En la figura 5.21 se visualiza una muestra de la imagen de entrenamiento perteneciente a este banco de imágenes.



Figura 5.21 Muestra de la imagen de entrenamiento y las correspondientes a la etapa de evaluación del banco de imágenes Faces94.

En la tabla 5.5 se muestran los resultados obtenidos con el algoritmo de pre-procesamiento propuesto (FRRM) y los diferentes clasificadores. Se observa que el porcentaje más alto de exactitud es del 96.08% con el clasificador SVM.

Tabla 5.5 Resultados de clasificación del banco de imágenes Faces94 en combinación con el algoritmo de pre-procesamiento propuesto FRRM.

	Exactitud	Tasa de error	Sensibilidad	Especificidad	Precisión
SVM	96.08	3.93	96.07	99.98	96.53
Redes bayesianas	91.51	8.52	91.48	99.94	92.91
MLP	91.71	8.32	91.68	99.94	92.93
Random forest	94.11	5.91	94.09	99.96	94.84

En la tabla 5.6 se muestran los resultados de exactitud de todos los algoritmos de pre-procesamiento. En ellos se observa, con letras en verde, los porcentajes más altos alcanzados por los algoritmos de pre-procesamiento, siendo el mejor resultado 96.08, logrado con la aplicación del algoritmo propuesto FRRM y el clasificador SVM. En segundo lugar, se encuentra la combinación de los algoritmos retina + retinex multi-escala con un 95.98%.

Tabla 5.6 Resultados de exactitud del banco de imágenes Faces94 respecto a los algoritmos de pre-procesamiento.

	FRRM	Retina	Retinex m.	EH	CLAHE	EH + retina	EH + retinex m.	CLAHE + retina	CLAHE + retinex m.	Retina+ retinex m.
SVM	96.08	95.75	95.36	95.19	94.4	95.59	94.9	95.85	93.55	95.98
Redes bayesianas	91.51	92	87.76	89.7	87.39	90.92	89.66	89.07	86.7	91.8
MLP	91.71	92.26	91.9	92.1	90.92	93.19	91.71	83.9	87.92	90.68
Random forest	94.11	93.98	91.54	92.53	91.77	92.79	91.44	93.32	90.39	94.24

En la figura 5.22 se muestra la gráfica comparativa de los valores más altos de exactitud respecto de la tabla 5.6. En la cual se puede observar que de manera semejante al banco de imágenes anterior, en este experimento, el clasificador SVM logra los mejores porcentajes con todos los algoritmos de pre-procesamiento.

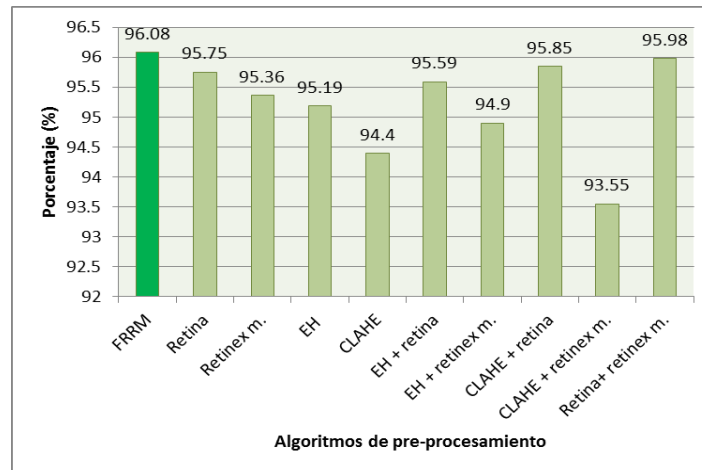
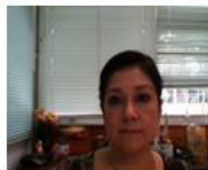


Figura 5.22 Comparativa de exactitud de los algoritmos de pre-procesamiento respecto al banco de imágenes Faces94.

Banco de imágenes propio

Se utilizaron 20 imágenes de entrenamiento correspondientes a una por aspirante (clase), así como 30 capturas obtenidas del video para la predicción, teniendo un total de 600 imágenes para la etapa de prueba. En la figura 5.23 se visualiza una muestra de la imagen de entrenamiento y evaluación de un aspirante perteneciente a este banco de imágenes. Como se puede ver, la imagen de entrenamiento contiene un rostro de frente, mientras que las imágenes correspondientes a la evaluación, presentan condiciones reales como movimiento, desenfoque, rostro no completo o con cierta rotación, entre otros aspectos.

Imagen de entrenamiento



Imágenes para clasificación



Figura 5.23 Muestra de la imagen de entrenamiento y las correspondientes a la etapa de evaluación del banco de imágenes propio.

En la tabla 5.7 se muestran los resultados obtenidos con el algoritmo de pre-procesamiento propuesto (FRRM) y los diferentes clasificadores. Se observa que el porcentaje más alto de exactitud es 95.14% obtenido con el clasificador de redes bayesianas. En esta prueba, SVM queda en segundo lugar, con un 92.88%. Muy abajo se encuentran los valores logrados por los otros 2 algoritmos de clasificación considerados.

Tabla 5.7 Resultados de clasificación en combinación con el algoritmo propuesto FRRM respecto al banco de imágenes propio, expresado en porcentajes.

	Exactitud	Tasa de error	Sensibilidad	Especificidad	Precisión
SVM	92.88	7.39	92.61	99.6	93.65
Redes bayesianas	95.14	4.78	95.22	99.73	95.5
MLP	76.91	22.73	77.27	98.49	82.15
Random forest	84.03	15.89	84.11	99.02	84.95

En la tabla 5.8 se muestran los resultados, con la métrica de exactitud, de todos los algoritmos de pre-preprocesamiento. En la cual se observa con letras verdes los porcentajes más altos alcanzados por algoritmo de pre-procesamiento; siendo el de mayor valor el 95.14% con el algoritmo propuesto y redes Bayesianas. A diferencia de las pruebas anteriores, en estos resultados el clasificador de redes Bayesianas sobresale con el mayor porcentaje. No obstante, en segundo lugar se encuentra nuevamente el algoritmo de pre-procesamiento propuesto con el clasificador SVM, siendo este resultado más alto que los obtenidos con los otros métodos. Es importante hacer notar que el método de retina queda en segundo lugar con un 92.19%, concordando con los resultados encontrados en la literatura, que lo reportan como bueno.

Tabla 5.8 Resultados de exactitud del banco de imágenes propio respecto a los algoritmos de pre-procesamiento.

	FRRM	Retina	Retinex m.	EH	CLAHE	EH + retina	EH + retinex m.	CLAHE + retina	CLAHE + retinex m.	Retina+ retinex m.
SVM	92.88	92.19	81.77	86.28	81.42	88.72	84.55	83.85	85.42	88.19
Redes bayesianas	95.14	85.24	81.25	80.21	71.53	82.29	83.33	71.7	86.11	88.37
MLP	76.91	72.92	76.22	90.45	66.49	71.18	79.51	77.84	74.31	83.85
Random forest	84.03	83.33	78.47	73.61	76.39	77.08	83.51	77.43	83.16	79.69

En la figura 5.24 se muestra la gráfica comparativa de los valores más altos de exactitud respecto de la tabla 5.8.

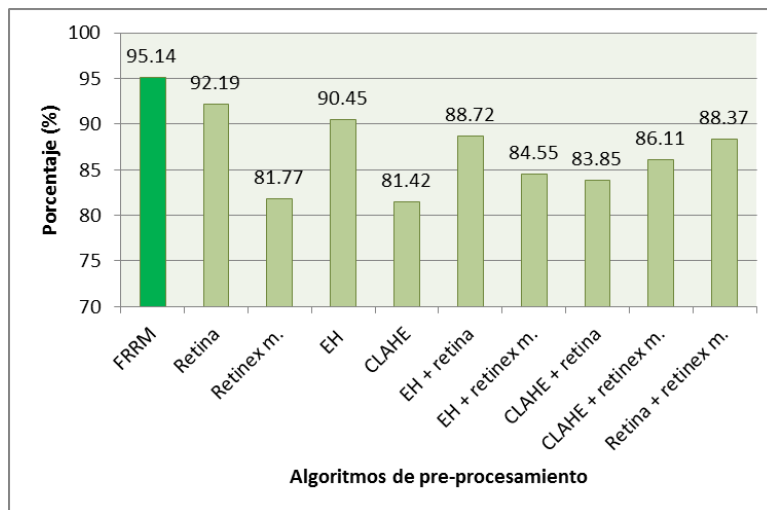


Figura 5.24 Comparativa de exactitud de los algoritmos de pre-procesamiento respecto a al banco de imágenes propio.

Comentarios finales de las pruebas de verificación de identidad

Como se muestra en la figura 5.25, la combinación del algoritmo de ecualización del histograma con el de retina logran el mejor porcentaje en el banco de imágenes FEI Face Database y en segundo lugar se encuentra el algoritmo propuesto (FRRM) teniendo una diferencia de 0.1%. Para el banco de imágenes Faces 94 el porcentaje más alto se alcanza con el algoritmo propuesto (FRRM) con un porcentaje del 96.08%, seguido de la combinación de algoritmos retina + retinex multi-escala, con una diferencia de 0.1%. Para el banco de imágenes propio el mejor rendimiento se alcanza con el algoritmo propuesto (FRRM) seguido del algoritmo de retina con una diferencia de 2.95%.

Como se puede apreciar en la gráfica, el método propuesto ayuda a alcanzar buenos porcentajes de clasificación comparado a los otros algoritmos de pre-procesamiento reportados en el estado del arte.

Es importante hacer notar que el algoritmo de pre-procesamiento propuesto (FRRM), mejora la calidad de las imágenes aún con cambios drásticos en la intensidad luminosa. No importando el rango dinámico presente en la imagen (con mucha o poca luz). También demuestra robustez ante los diferentes tonos de piel de la persona o presencia de algunas sombras en el rostro.

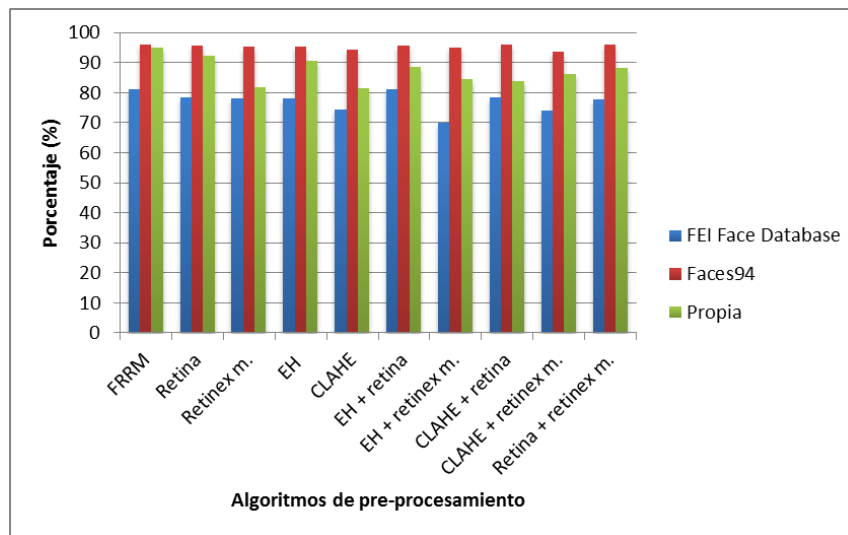


Figura 5.25 Comparación de resultados de clasificación respecto a los bancos de imágenes.

Tiempo

El coste computacional en el proceso de verificación de identidad tiene gran importancia en este proyecto, debido a que el sistema debe proporcionar una respuesta cercana al tiempo real por ser un sistema de exámenes en línea, los resultados se muestran a continuación.

En la tabla 5.9 se muestra el tiempo (en segundos) de la etapa de entrenamiento con los tres bancos de imágenes respecto a los algoritmos de pre-procesamiento. En la figura 5.26 se muestra la gráfica correspondiente a la tabla 5.9, en donde se puede observar que el algoritmo propuesto (barra en morado) se encuentra en un punto medio, concluyendo que a pesar de no ser el más rápido tiene tiempos aceptables. El menor tiempo lo logra el algoritmo de ecualización del histograma (Barra En verde) y el mayor coste es la combinación de los algoritmos de retina con retinex multi-escala (barra en rosa).

Tabla 5.9 Resultados de los tiempos de entrenamiento expresados en segundos.

	FRRM	Retina	Retinex m.	EH	CLAHE	EH + retina	EH + retinex m.	CLAHE + retina	CLAHE + retinex m.	Retina+ retinex m.
Banco de imágenes propio	65.81	70.44	40.47	38.73	40.14	74.29	39.43	72.03	40.94	75.89
FEI Face Database	561.60	588.80	445.60	283.20	256.81	599.34	286.03	599.21	289.13	606.34
Faces94	364.17	372.16	151.72	141.81	149.58	436.57	157.68	384.23	180.64	399.48

A pesar de que Faces 94 tiene el mayor número de imágenes, se puede observar en la figura 5.26, que el mayor tiempo lo tiene el banco de imágenes *FEI Face Database*. Esto obedece a que la resolución de las imágenes en *Faces94* es de 180 x 200. Mientras que el banco de imágenes FEI maneja resoluciones de 640 x 480. La diferencia en

tamaño tiene como consecuencia que los algoritmos de pre-procesamiento demoren más en procesar la imagen.

La diferencia de tiempos de entrenamiento entre el algoritmo propuesto (FRRM) y el de retina es de 4.63 segundos más. Mientras que la diferencia, respecto a la combinación del algoritmo retina + retinex multi-escala, es de 10.08 segundos menos, colocando al algoritmo propuesto como una solución eficiente y relativamente rápida.

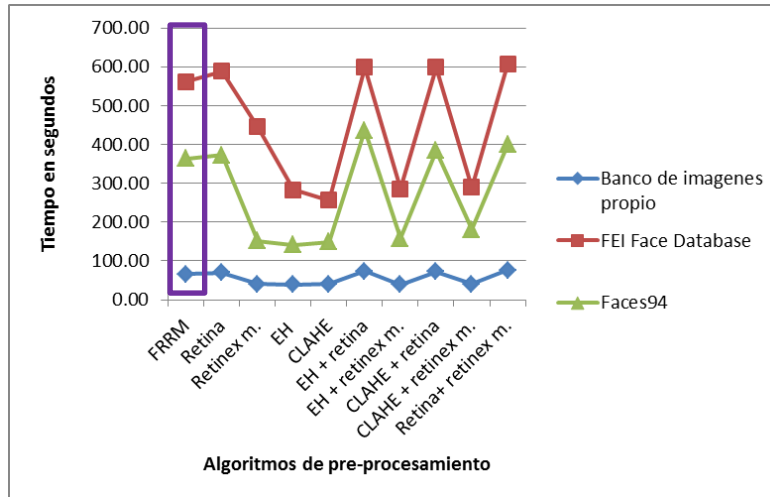


Figura 5.26 Grafica de tiempos de entrenamiento expresados en segundos.

Con respecto a la etapa de evaluación, en la tabla 5.10 se muestra el tiempo promedio en segundos, que tarda el sistema en verificar la identidad de un aspirante para los bancos de imágenes respecto a los algoritmos de pre-procesamiento. En la figura 5.27 se muestra la gráfica correspondiente a la tabla 5.10, en donde se puede observar que el tiempo en esta etapa, del algoritmo propuesto se encuentra por debajo de 0.3 segundos (barra en morado), los mejores tiempos se obtienen con el algoritmo de ecualización del histograma (barra en verde) y los más altos con la combinación de los algoritmos de retina y retinex multi-escala (tono rosa).

Tabla 5.10 Resultados de los tiempos de monitoreo expresados en segundos.

	FRRM	Retina	Retinex m.	EH	CLAHE	EH + retina	EH + retinex m.	CLAHE + retina	CLAHE + retinex m.	Retina+ retinex m.
Banco de imágenes propio	0.26	0.27	0.15	0.13	0.14	0.29	0.14	0.28	0.15	0.30
FEI Face Database	0.27	0.30	0.13	0.12	0.13	0.27	0.13	0.27	0.13	0.28
Faces94	0.22	0.24	0.09	0.08	0.08	0.25	0.09	0.23	0.10	0.24

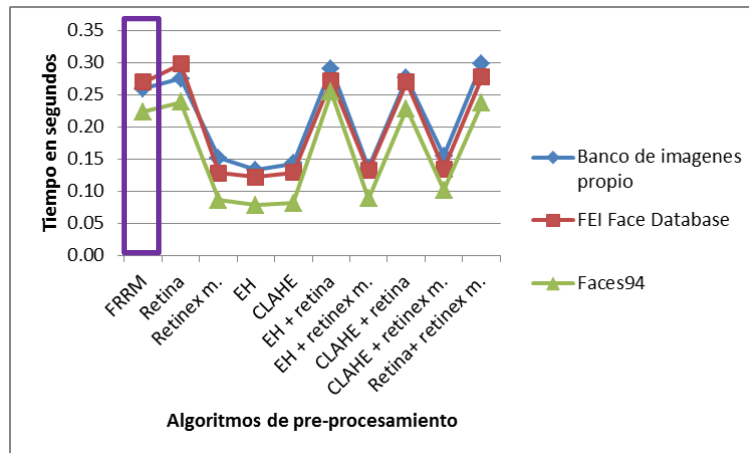


Figura 5.27 Grafica de tiempos de monitoreo promedio por imagen expresados en segundos.

5.4 Comentarios finales

Después de realizar las diferentes pruebas se concluye lo siguiente: tomando como referencia la métrica SSIM, el algoritmo de pre-procesamiento que entrega mejores resultados es la combinación de la ecualización del histograma en combinación con el modelo de retina alcanzando un promedio del 98% de similitud; sin embargo, de manera visual los mejores resultados se obtienen en el modelo propuesto, debido a que, lo que se busca no es una imagen homogénea si no que exista un realce en las zonas de interés del rostro (ojos, nariz y boca) para lograr mayor exactitud en la localización de los puntos faciales.

Estos resultados son corroborados en la etapa de clasificación, donde el algoritmo propuesto FRRM logra resultados superiores (en la métrica de exactitud y en todas las demás) en el banco de imágenes propio con 95.14% y Faces94 con 96.08%, superando a los otros algoritmos de pre-procesamiento. En el banco de imágenes *FEI Face Database* el algoritmo con mayor porcentaje de exactitud es la combinación de Ecualización del histograma con retina alcanzando 81.22%, siendo 0.1% superior al algoritmo propuesto que alcanza 81.12% de exactitud.

En cuanto a tiempos, el algoritmo propuesto reporta tiempo aceptables al no superar los 0.3s en la verificación de identidad situándose en un punto medio de los otros algoritmos de pre-procesamiento.

CAPÍTULO 6

Conclusiones

6.1 Conclusiones

En este trabajo se presentó la metodología usada para llevar a cabo las tareas de pre-procesamiento de imágenes y verificación de identidad, así como la experimentación y los resultados.

Se presentan 5 algoritmos de pre-procesamiento que son: ecualización del histograma (EH), ecualización adaptativa (CLAHE), modelo de retina, retinex multi-escala y el algoritmo propuesto (FRRM). Así como, cuatro combinaciones entre ellos que son: EH + retina, EH + retinex multi-escala, CLAHE + retina y CLAHE + retinex multi-escala. Con los resultados obtenidos se observa que el algoritmo propuesto (FRRM) tiene resultados superiores a la EH, algoritmo clásico más utilizado en la literatura para abordar el problema de la iluminación variable en imágenes faciales, así como el algoritmo de retina que es de las técnicas más recientes y con mejores resultados reportados en el estado del arte, superándolos incluso en tiempo, lo cual es importante en sistemas con respuesta en tiempo real.

Para la verificación de identidad se compara el rendimiento de cuatro clasificadores que son: máquinas de vector soporte (SMV), perceptrón multicapa, redes bayesianas y random forest; los cuales son evaluados considerando: exactitud, precisión, especificidad, sensibilidad y tasa de error. Después de los experimentos se concluye que el algoritmo de pre-procesamiento propuesto en combinación con el clasificador SVM alcanza el mejor rendimiento en comparación con las otras técnicas y clasificadores; siendo las máquinas de vector soporte uno de los algoritmos más utilizados dentro del estado del arte.

El clasificador que se utiliza en el funcionamiento final del sistema es máquinas de vector soporte (SVM), debido a que en la experimentación el comportamiento del vector de características con el clasificador tiene buen desempeño. Los resultados del sistema en porcentaje de exactitud y tiempo en la etapa de verificación de identidad muestran buenos resultados para la confiabilidad del sistema, siendo así la combinación del algoritmo propuesto (FRRM) con el clasificador de SVM, la combinación óptima en comparación con el resto de los algoritmos de pre-procesamiento y clasificadores.

En conclusión, el problema del reconocimiento facial en ambientes naturales es un tema complejo en el cual se sigue trabajando; ya que como se puede observar en los resultados reportados se muestra que un buen pre-procesamiento de las imágenes faciales es fundamental para que la detección de los puntos característicos que conforman la máscara sea lo más apegada posible a las zonas de interés (ojos, boca y nariz) y con ello lograr un buen reconocimiento facial.

6.2 Objetivos logrados

En la tabla 6.1 se muestra una breve descripción de cómo se abordó cada objetivo específico.

Tabla 6.1 Avance de logros por objetivo específico.

Objetivo específico	Solución del objetivo
<p>Revisar en el estado del arte las técnicas utilizadas para reconocimiento facial con limitaciones de iluminación</p>	<p>Se analizaron 60 y se reportan los 27 artículos más relacionados, en las temáticas de procesamiento de iluminación, estudios de reconocimiento facial, etc.</p> <p>Se seleccionaron e implementaron 4 algoritmos utilizados en la literatura para abordar las limitaciones de iluminación en el pre-procesamiento de imágenes faciales los cuales son:</p> <ul style="list-style-type: none"> • Ecuilización del histograma (EH) • Ecuilización adaptativa (CLAHE) • Retina • Retinex multi-escala. <p>Se realizaron pruebas de combinaciones de los algoritmos anteriores las cuales son:</p> <ul style="list-style-type: none"> • EH + retina • EH + retinex multi-escala • CLAHE + retina • CLAHE + retinex multi-escala <p>Derivado del análisis de las técnicas anteriores, se desarrolló una propuesta que fusiona los algoritmos de retina y retinex multi-escala.</p>
<p>Analizar los componentes del rostro que describen mejor a una persona en un ambiente natural.</p>	<p>A partir de la máscara de puntos de la librería Dlib [42], se analizaron y evaluaron de manera experimental diferentes combinaciones de distancias faciales para seleccionar el menor número de las mismas y que además, dichas distancias tuvieran la característica de ser discriminativas e invariantes (lo más posible) a cambios geométricos como la escala, rotación y traslación.</p>
<p>Diseñar e implementar un sistema que realice el reconocimiento con diferentes condiciones de intensidad luminosa.</p>	<p>Se desarrolló un sistema que, a través de la técnica de pre-procesamiento propuesta, ayuda a mejorar el rendimiento del clasificador en la verificación de identidad del aspirante, reduciendo el margen de error causado por las diferentes condiciones de intensidad luminosa.</p>
<p>Identificar a la persona en un ambiente no controlado.</p>	<p>El sistema realiza diferentes capturas de imágenes del aspirante hasta encontrar una que contenga la cara lo suficientemente de frente para poder detectar todos los componentes del rostro (ojos nariz y boca) y así poder iniciar con el análisis automático de la imagen y llevar a cabo la verificación de identidad del aspirante.</p>
<p>Proporcionar una respuesta cercana al tiempo real</p>	<p>El sistema, durante el monitoreo del aspirante, proporciona una respuesta de verificación de identidad por debajo de los 0.3 segundos.</p>

6.3 Aportaciones

Las aportaciones que se obtuvieron en la realización de este proyecto de investigación son:

1. Algoritmo de pre-procesamiento de imágenes.
Se propone un algoritmo de pre-procesamiento para imágenes con problemas de iluminación, que comprende la fusión entre el modelo de retina y el algoritmo de Retinex multi-escala, esto para lograr la homogeneidad en las imágenes que proporciona el modelo de retina y resaltar los puntos característicos del rostro mediante el contraste que proporciona el algoritmo retinex.
Específicamente el algoritmo fusiona la capa de foto receptores, utilizada en el modelo de retina y del algoritmo retinex multi-escala las envolventes gaussianas de diferentes anchos para asegurar una compresión del rango dinámico y un rendimiento en el ajuste tonal adecuado.
2. Verificación de identidad
Se hace una comparativa de diferentes clasificadores utilizados en la literatura. Se concluye que el uso del clasificador máquinas de vector soporte tiene una alta precisión en la verificación de identidad.
3. Respuesta en tiempo real
Después de realizar las pruebas de verificación de identidad con los diferentes bancos de imágenes se comprueba que el algoritmo propuesto tiene un tiempo cercano al real logrando una ventaja en el sistema al momento de monitorear a los aspirantes.
4. Sistema de monitoreo para verificación de identidad para valuación en línea
Se propone un novedoso sistema en donde con una interfaz sencilla para el usuario, se puede realizar un entrenamiento rápido con las imágenes de los aspirantes y un monitoreo sencillo que pueda reducir el tiempo de monitoreo manual por parte del encargado de monitorear el examen.

6.4 Trabajo futuro

Como trabajos futuros se tiene el siguiente punto:

- Integración del modelo de pre-procesamiento de imágenes y verificación de identidad con el sistema de evaluación en línea del centro nacional de investigación y desarrollo tecnológico (CENIDET). Actividad no contemplada en los alcances de este proyecto.

6.5 Productos académicos adicionales

- Publicación del artículo titulado “Comparación de Algoritmos de Clasificación con Reducción de Variables para Reconocimiento Facial”, en el congreso CIINDET Marzo 2018, con ISBN 978-607-95255-8-3, pág. 191-195. (anexo B figura B.1, figura B.2)
- Presentación de poster y ponencia con el tema “Reconocimiento automático del rostro para verificación de identidad para evaluación en línea”, en la escuela de Inteligencia Artificial y Robótica 2018, realizada en la universidad tecnológica Emiliano Zapata (UTEZ) los días 25 y 26 de Octubre. (anexo B figura B.3)
- Participación con el póster “Reconocimiento automático del rostro para verificación de identidad para evaluación en línea”, en la 1A jornada de Ciencia y Tecnología aplicada, realizada en el Centro Nacional de Investigación y Desarrollo Tecnológico, los días 15 y 16 de noviembre de 2018. (anexo B figura B.4)
- Participación y asistencia al evento de Innovación con el tema “Reconocimiento automático del rostro para verificación de identidad para evaluación en línea”, que se celebró en el Congreso Internacional de Ingeniería Mecatrónica, Electrónica y Automática (ICMEAE) el 28 y 29 de Noviembre de 2018. (anexo B figura B.5)
- Presentación del poster titulado “Suplantación de identidad en imágenes faciales” en la Segunda Jornada de Ciencia y Tecnología aplicada, a realizarse en el Centro Nacional de Investigación y Desarrollo Tecnológico, los días 4 y 5 de Abril de 2019. (anexo B figura B.6)
- Presentación poster y ponencia con tema “Reconocimiento automático del rostro para verificación de identidad para evaluación en línea”, en el primer coloquio de inteligencia artificial, realizada en el Instituto Tecnológico de Iguala (ITI) el día 9 de mayo del 2019. (anexo B figura B.7)

Anexos

A. Manual de instalación (OpenCV [41], Dlib [42] y Qt [45])

Para el desarrollo del sistema se utilizaron las siguientes herramientas.

- Visual Studio community 2015
- CMAKE 3.12.1 [48]: CMake es una familia de herramientas multiplataforma de código abierto diseñada para compilar, probar y empaquetar software. CMake se utiliza para controlar el proceso de compilación de software mediante la plataforma simple y los archivos de configuración independientes del compilador, y generar *makefiles* y espacios de trabajo nativos que se pueden usar en el entorno de compilación de su elección.
- OpenCV 3.4.0 [41]: OpenCV es una biblioteca libre de visión artificial originalmente desarrollada por Intel.
- Dlib 19.15 [42]: Dlib es una biblioteca de software multiplataforma de uso general escrita en el lenguaje de programación C++.
- Qt 4.8.1 [45]: Para la creación de la interfaz gráfica

Procedimiento de instalación

1. Se descarga e instala visual studio community 2015 de Microsoft.

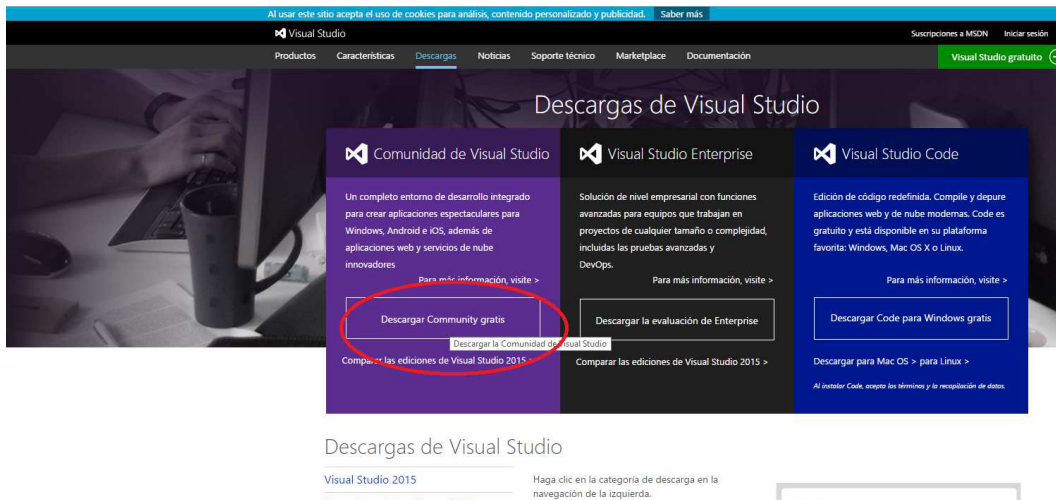


Figura A.1 Pantalla “Descargas de Visual Studio”.

2. Se descarga OpenCV y se descomprime en la dirección de su elección, en este caso “C:”. Se descargan las librerías adicionales “opencv_contrib” que corresponda a la versión que se utiliza de opencv, para esta instalación la versión es 3.4.0.

3. Una vez que se descomprime en la carpeta de su elección se agrega la dirección a las variables de sistema, para ello se va a sistema en panel de control, configuración avanzada del sistema.

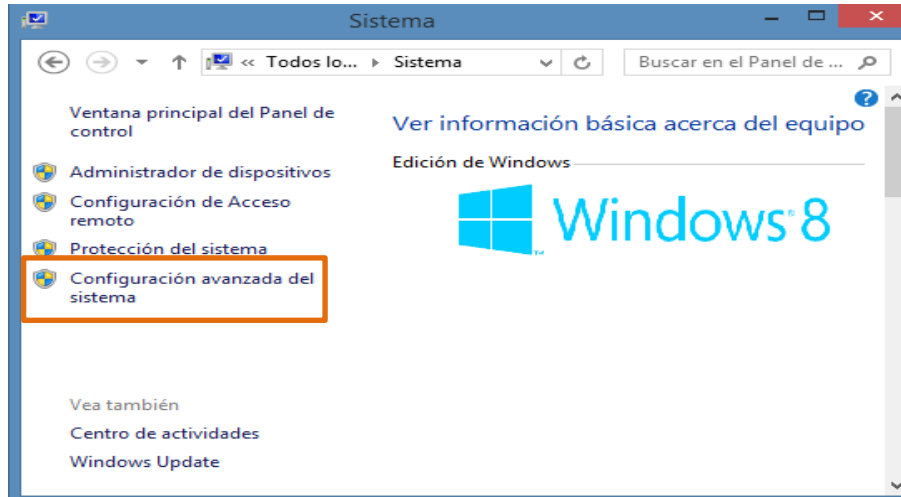


Figura A.2 Pantalla de configuración del sistema.

- a. En propiedades del sistema seleccionamos variables de entorno.

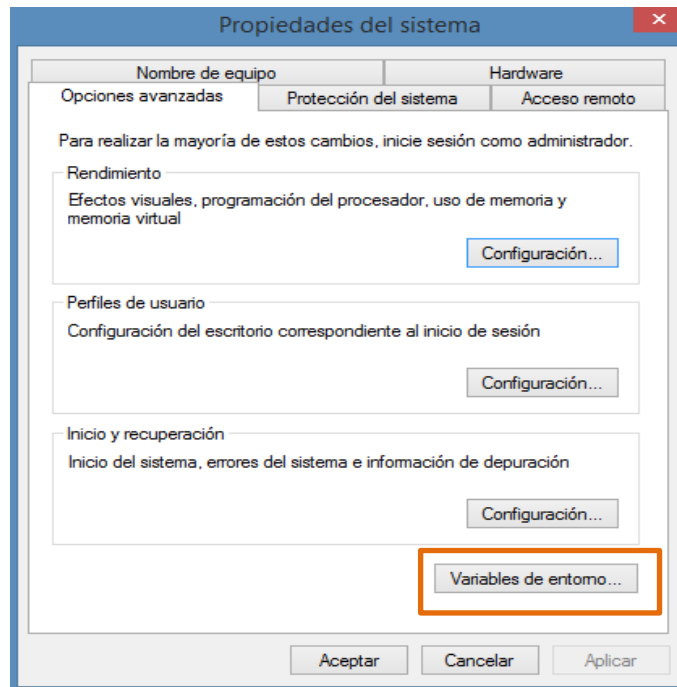


Figura A.3 Pantalla de propiedades del sistema.

- b. En la variable Path se agrega la dirección de OpenCV.

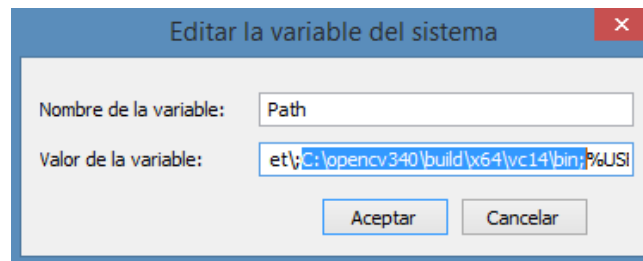


Figura A.4 Ventana de edición de variables del sistema.

4. Se descarga Dlib y se descomprime dentro de la carpeta de OpenCV.
5. Se descarga e instala CMAKE, una vez instalado se crea la compilación de OpenCV seleccionando las carpetas en donde está contenido OpenCV

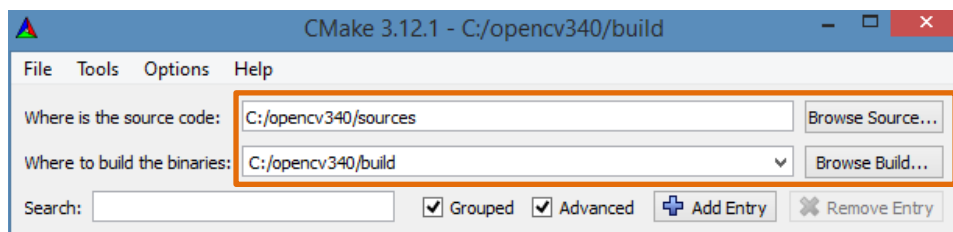


Figura A.5 Selección de carpetas donde está contenida la carpeta OpenCV.

- a. Se selecciona configurar, en esta opción se selecciona la versión de visual studio para la que se generara la compilación, en este caso Visual Studio 14 2015 Win64

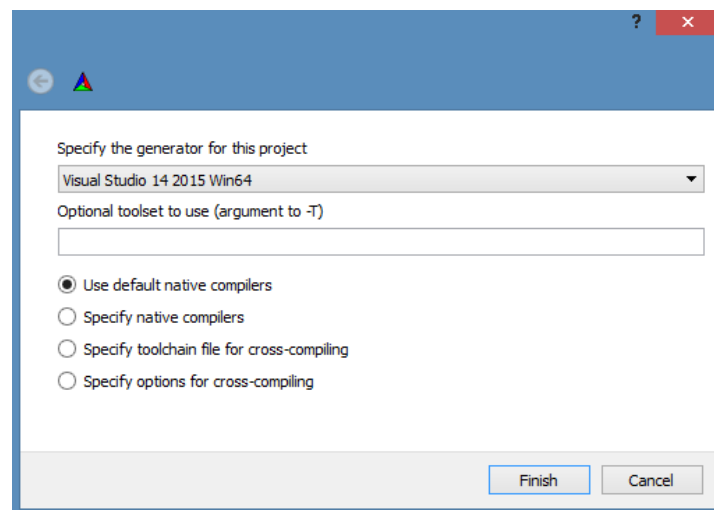


Figura A.6 Selección de la versión de visual studio que se desea configurar.

- b. Se selecciona finalizar y enseguida configurar

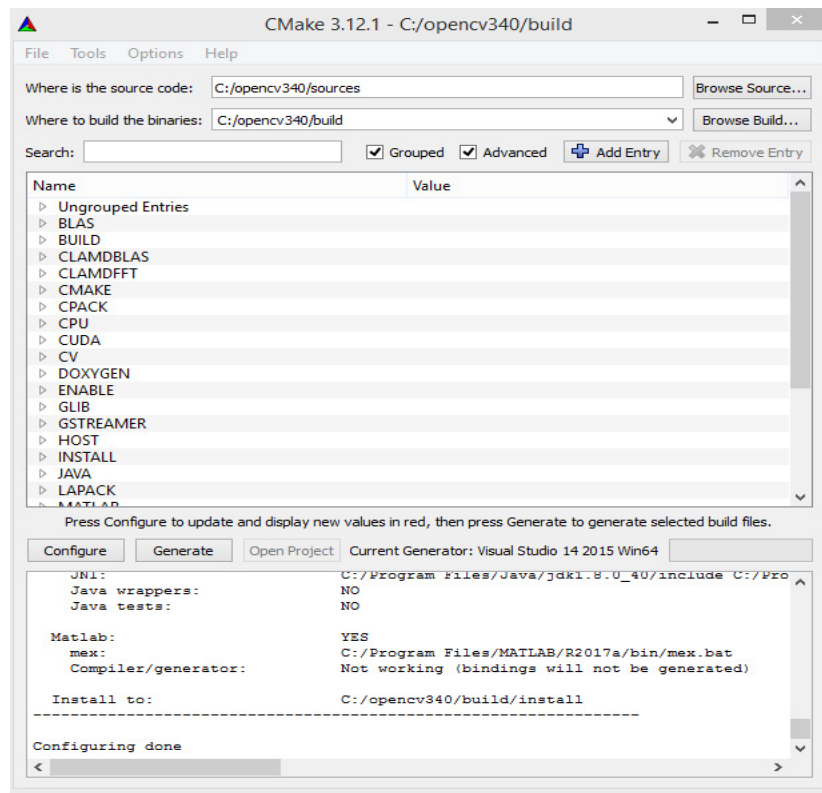


Figura A.7 Proceso de configuración.

- c. En “OPENCV-OPENCV_EXTRA_MODULES_PATH” seleccionamos la dirección que contiene opencv_contrib

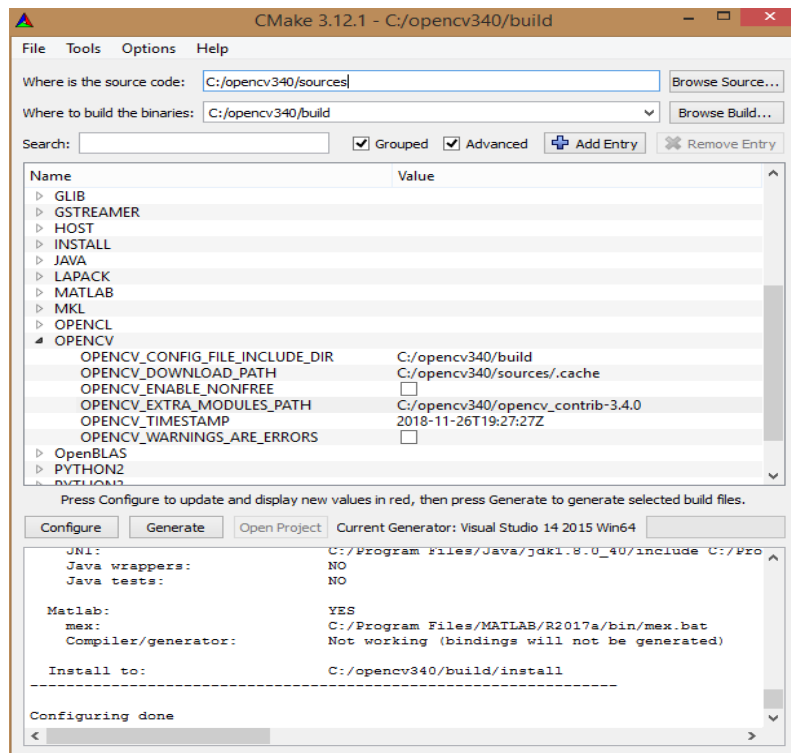


Figura A.8 Selección de los módulos adicionales.

- d. Se vuelve a seleccionar configurar y enseguida generar

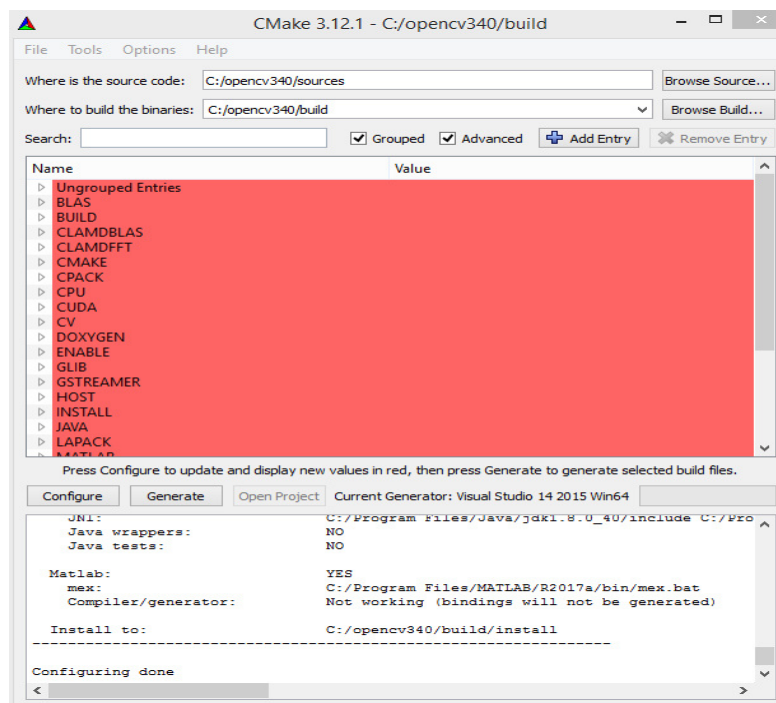


Figura A.9 Generación de la configuración.

- Una vez que se generaron los archivos para la configuración de OpenCV se ejecuta el archivo OpenCV.sln

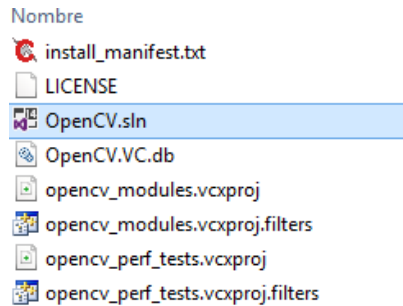


Figura A.10 Pantalla de selección del archivo .sln

- Una vez que se abre el archivo .sln se compila la solución para generar el archivo Debug y Release
- Ahora se repite lo mismo pero para compilar Dlib

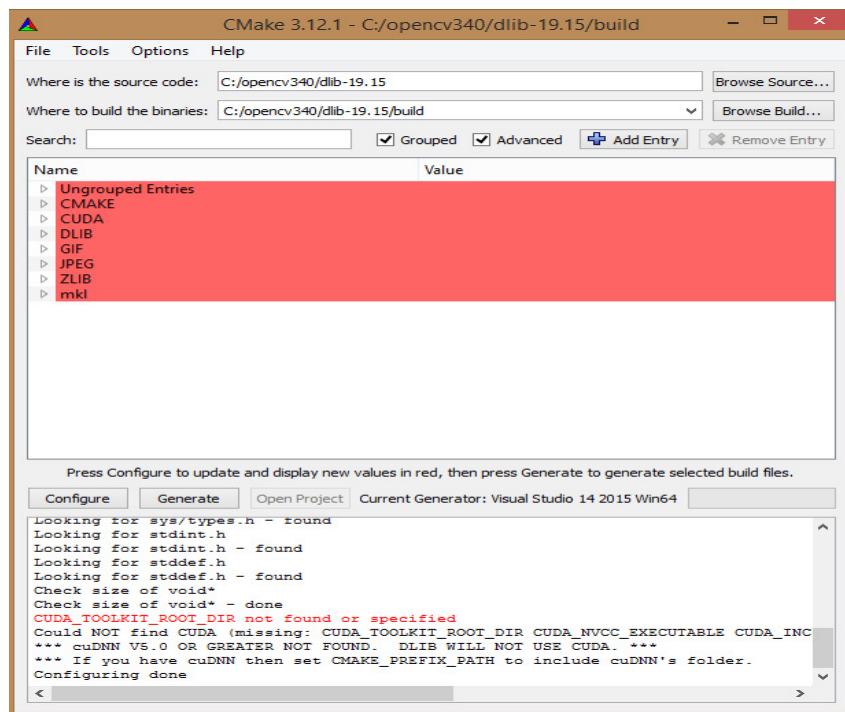


Figura A.11 Pantalla de configuración para Dlib.

- Al igual que en OpenCV se ejecuta el archivo .sln y se compila

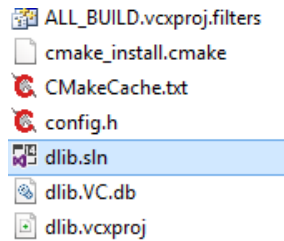


Figura A.12 Selección del archivo de configuración .sln para Dlib.

10. Una vez que se compilaron las dos soluciones se configura visual studio
 - a. Seleccionamos Directorios de VC++
 - i. En el apartado Directorios de archivos ejecutables agregamos la liga de instalación de OpenCV (C:\opencv340_2\build\install\x64\vc14\bin)

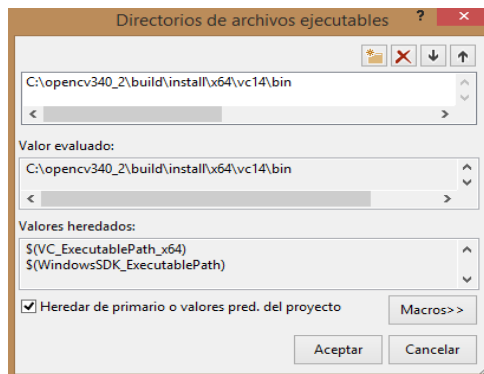


Figura A.13 Configurar visual studio en sección VC++

- ii. En el apartado Directorios de archivos de inclusión agregamos la dirección de Dlib (C:\opencv340_2\dlib-19.4)

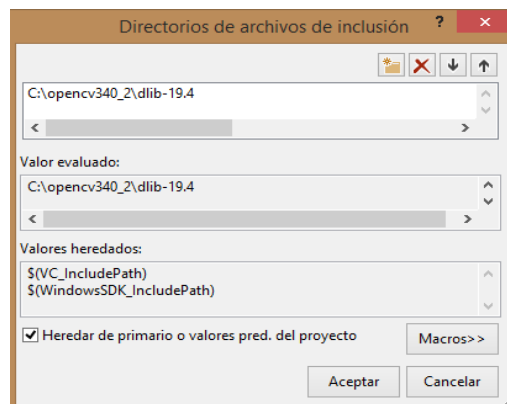


Figura A.14 Configurar visual studio en sección directorio de archivos de inclusión.

- iii. En directorio de archivos de biblioteca agregamos la configuración de opencv y Dlib

(C:\opencv340_2\build\install\x64\vc14\lib
C:\opencv340_2\dlib-19.4\build\Release)

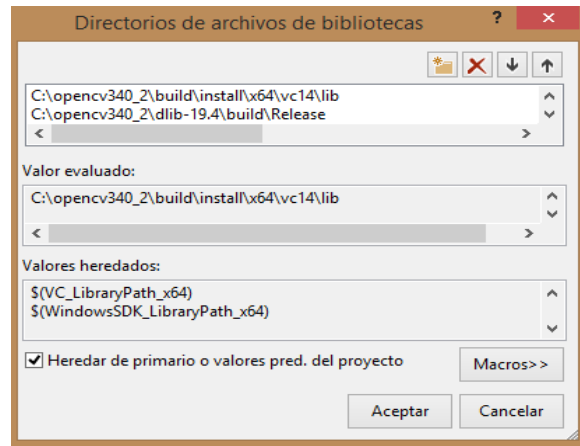


Figura A.15 Configurar visual studio en sección directorios de archivos de biblioteca.

- b. Nos vamos a la sección de Vinculador
 - i. En el apartado Directorios de bibliotecas adicionales agregamos la dirección de compilación Dlib (C:\opencv340_2\dlib-19.4\build\Release)

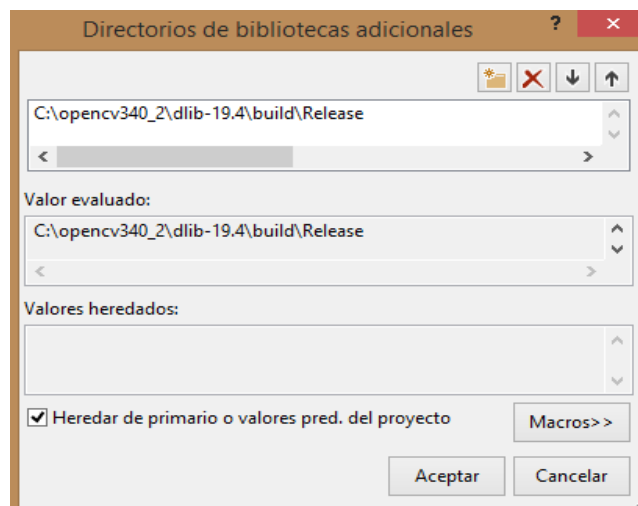


Figura A.16 Configurar visual studio en sección directorio de bibliotecas adicionales

- ii. En el apartado Entradas – Dependencias adicionales agregamos todas las librerías que se utilizaran
(dlib.lib, opencv_aruco340.lib, opencv_bgsegm340.lib,
opencv_bioinspired340.lib
opencv_calib3d340.lib, opencv_ccalib340.lib, opencv_core340.lib,
opencv_datasets340.lib)

opencv_dnn340.lib, opencv_dpm340.lib, opencv_face340.lib,
opencv_features2d340.lib
opencv_flann340.lib, opencv_fuzzy340.lib, opencv_highgui340.lib
opencv_img_hash340.lib, opencv_imgcodecs340.lib,
opencv_imgproc340.lib, opencv_line_descriptor340.lib
opencv_ml340.lib, opencv_objdetect340.lib, opencv_optflow340.lib
opencv_phase_unwrapping340.lib, opencv_photo340.lib,
opencv_plot340.lib
opencv_reg340.lib, opencv_rgbd340.lib, opencv_saliency340.lib,
opencv_shape340.lib
opencv_stereo340.lib, opencv_stitching340.lib,
opencv_structured_light340.lib, opencv_superres340.lib,
opencv_surface_matching340.lib
opencv_text340.lib, opencv_tracking340.lib, opencv_video340.lib
opencv_videoio340.lib, opencv_videostab340.lib,
opencv_xfeatures2d340.lib
opencv_ximgproc340.lib, opencv_xphoto340.lib)

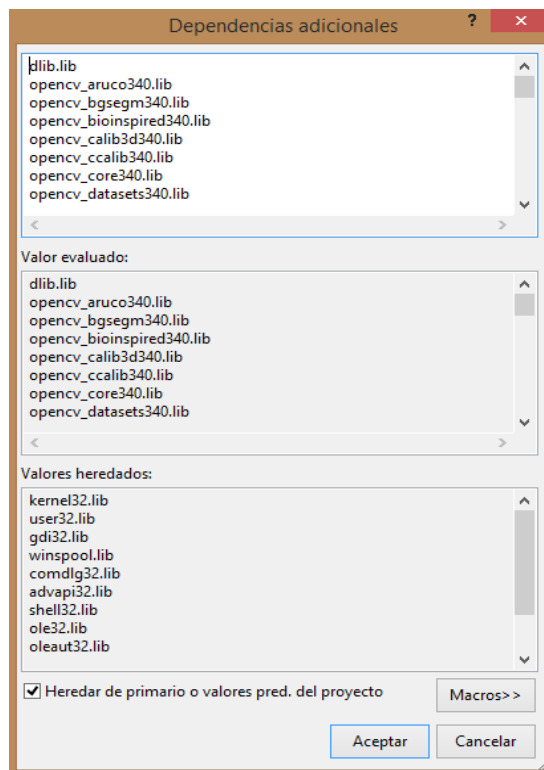


Figura A.17 Configurar visual studio en sección dependencias adicionales.

- c. Se crea un nuevo proyecto en visual studio para activar el menú C/C++. Dentro de este apartado agregamos la dirección a opencv (C:\opencv340_2\build\install\include)

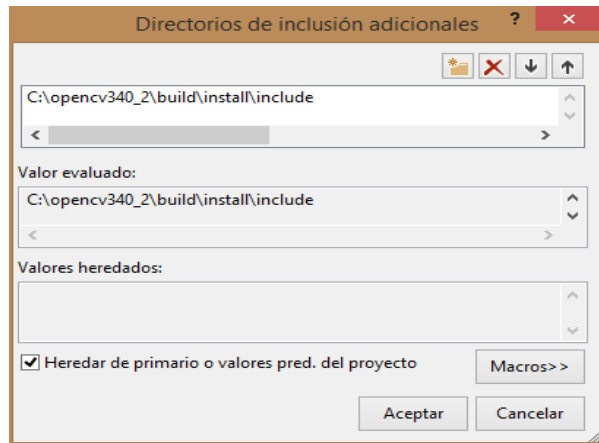


Figura A.18 Configurar visual studio en sección directorios de inclusión adicionales.

11. Se agregan las variables de QT en visual studio. C/C++->General->Directorios de inclusión adicionales (C:\Qt\5.12.1\msvc2015_64\include\QtANGLE)

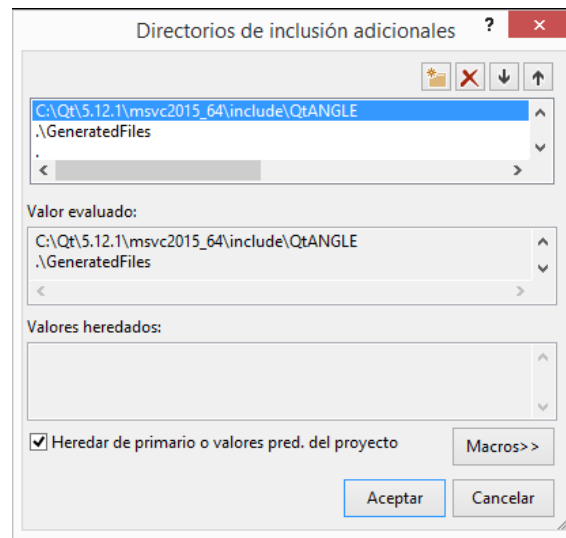


Figura A.19 Configurar visual studio en sección directorios de inclusión adicionales para Qt

12. Por último se agrega a las variables de sistema la dirección de la instalación de openCV

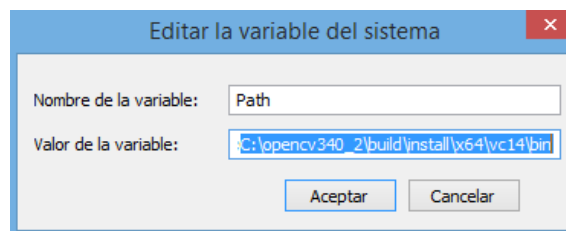


Figura A.20 Ventana de edición de variables del sistema.

B. Productos académicos adicionales



Figura B.1 Constancia por la publicación del artículo titulado “Comparación de algoritmos de clasificación con reducción de variables para reconocimiento facial” en el congreso CIINDET 2018.



Figura B.2 Constancia por la participación como congresista con el artículo titulado “Comparación de algoritmos de clasificación con reducción de variables para reconocimiento facial” en el congreso CIINDET 2018.



Figura B.3 Reconocimiento por la participación como ponente de la conferencia “Reconocimiento automático del rostro para verificación de identidad para evaluación en línea” en el marco del evento: Escuela de inteligencia artificial y robótica 2018, llevado a cabo en las instalaciones de la universidad tecnológica Emiliano Zapata.



Figura B.4 Constancia por la participación con el póster “Reconocimiento automático del rostro para verificación de identidad para evaluación en línea”, en la 1A jornada de Ciencia y Tecnología aplicada.



Figura B.5 Reconocimiento por la participación como concursante en la categoría innovación en el Advanced Robotics and Drone Competition que se celebró en el Congreso Internacional de Ingeniería Mecatrónica, Electrónica y Automática (ICMEAE).



Figura B.6 Constancia por la participación con el poster titulado “Suplantación de identidad en imágenes faciales” en la Segunda Jornada de Ciencia y Tecnología aplicada.



Figura B.7 Reconocimiento por la participación como ponente de la conferencia denominada “Reconocimiento automático del rostro para verificación de identidad para evaluación en línea” en el marco del 1er coloquio “Artificial Intelligence” realizada en el Instituto Tecnológico de Iguala (ITI).

Referencias

- [1] R. Subban and D. P. Mankame, "Human face recognition biometric techniques: Analysis and review," *Advances in Intelligent Systems and Computing*, vol. 235, pp. 455–463, 2014.
- [2] L.-F. Zhou, Y.-W. Du, W.-S. Li, J.-X. Mi, and X. Luan, "Pose-Robust Face Recognition with Huffman-LBP Enhanced by Divide-and-Rule Strategy," *Pattern Recognition* 78, pp. 43–55, 2018.
- [3] Z. Tang, X. Wu, B. Fu, W. Chen, and H. Feng, "Fast face recognition based on fractal theory," *Applied Mathematics and Computation* 321, pp. 721–730, 2018.
- [4] Y. F. Yu, D. Q. Dai, C. X. Ren, and K. K. Huang, "Discriminative multi-scale sparse coding for single-sample face recognition with occlusion," *Pattern Recognition* 66, pp. 302–312, 2017.
- [5] M. Hassaballah and S. Aly, "Face recognition : challenges , achievements and future directions," *IET Computer Vision*, vol. 9(4), pp. 614–626, 2015.
- [6] P. Kaur and L. Clay, "A Survey on Face Recognition Techniques," *International Journal of Advanced Research in Computer Science*, vol. 8, no. 4, 2017.
- [7] P. Wagh, J. Chaudhari, R. Thakare, and S. Patil, "Attendance System based on Face Recognition using Eigen face and PCA Algorithms," *International Conference on Green Computing and Internet of Things (ICGCIoT)*, pp. 303–308, 2015.
- [8] L. Liu, P. Fieguth, G. Zhao, M. Pietikäinen, and D. Hu, "Extended Local Binary Patterns for Face Recognition," *Information Sciences* 358, pp. 56–72, 2016.
- [9] M. M. Ghazi and H. K. Ekenel, "A Comprehensive Analysis of Deep Learning Based Representation for Face Recognition," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 34–41, 2016.
- [10] M. A. Ochoa-villegas, J. A. Nolasco-flores, O. Barron-cano, and I. A. Kakadiaris, "Addressing the illumination challenge in two- dimensional face recognition : a survey," *IET Computer Vision*, vol. 9, pp. 978–992, 2015.
- [11] A. H. Sable, P. Sanjay, and N. Talbar, "A Novel Illumination invariant Face recognition method based on PCA and WPD using YCbCr color space," *Procedia Computer Science*, vol. 92, pp. 181–187, 2016.
- [12] A. Kumar and Y. Narain, "Evaluation of Face Recognition Methods in Unconstrained Environments," *Procedia Computer Science*, vol. 48, pp. 644–651, 2015.
- [13] L. Zhuang, T. Chan, A. Y. Yang, S. S. Sastry, and Y. Ma, "Sparse Illumination Learning and Transfer for Single-Sample Face Recognition with Image Corruption and Misalignment," *International Journal of Computer Vision*, vol. 114, pp. 272–287, 2015.
- [14] N. McLaughlin, J. Ming, and D. Crookes, "Largest Matching Areas for Illumination and Occlusion Robust Face Recognition," *IEEE Transactions on Cybernetics*, vol. 47, no. 3, pp. 796–808, 2016.
- [15] S. Karamizadeh, S. M. Abdullah, and E. Randjbaran, "Face Recognition by Implying Illumination Techniques – A Review Paper," *Journal of Science and Engineering*, vol. 6, no. 01, pp. 001–007, 2015.
- [16] J. H. Shah, M. Sharif, M. Raza, and M. Murtaza, "Robust Face Recognition Technique under Varying Illumination," *Journal of Applied Research and Technology*, vol. 13, no. 1, pp. 97–105, 2015.
- [17] S. Chaithra and K. Umapathy, "Effectual Face Recognition System for Uncontrolled Illumination," *International Journal of Data Mining Techniques and Applications*, vol. 5, no.

- 1, pp. 62–64, 2016.
- [18] A. Lumini, L. Nanni, and S. Brahmam, “Ensemble of texture descriptors and classifiers for face recognition,” *Applied Computing and Informatics*, vol. 13, no. 1, pp. 79–91, 2017.
- [19] F. Juefei-xu and M. Savvides, “Encoding and Decoding Local Binary Patterns for Harsh Face Illumination Normalization,” in *IEEE International Conference on Image Processing (ICIP)*, 2015, pp. 3220–3224.
- [20] A. Punnappurath *et al.*, “Face recognition across non-uniform motion blur, illumination, and pose,” *IEEE Transactions on image processing*, vol. 24, no. 7, pp. 2067–2082, 2015.
- [21] S. P. Mudunuri and S. Biswas, “Low Resolution Face Recognition Across Variations in Pose and Illumination,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 5, pp. 1034–1040, 2016.
- [22] F. Cao, H. Hu, J. Lu, J. Zhao, Z. Zhou, and J. Wu, “Pose and Illumination Variable Face Recognition via Sparse Representation and illumination dictionary,” *Knowledge-Based Systems*, vol. 107, pp. 117–128, 2016.
- [23] R. Shyam and Y. N. Singh, “Face Recognition using Augmented Local Binary Pattern and Bray Curtis Dissimilarity Metric,” in *Signal processing and integrated networks (SPIN)*, pp. 779–784, 2015.
- [24] H. Y. Patil, A. G. Kothari, and K. M. Bhurchandi, “Expression invariant face recognition using local binary patterns and contourlet transform,” *Optik - International Journal for Light and Electron Optics*, vol. 127, no. 5, pp. 2670–2678, 2016.
- [25] W. Huang and H. Yin, “Robust face recognition with structural binary gradient patterns,” *Pattern Recognition*, vol. 68, pp. 126–140, 2017.
- [26] L. A. Cament, F. J. Galdames, K. W. Bowyer, and C. A. Perez, “Face recognition under pose variation with local Gabor features enhanced by Active Shape and Statistical Models,” *Pattern Recognition*, vol. 48, no. 11, pp. 3371–3384, 2015.
- [27] H. Ouanan and B. Ouanan, M., & Aksasse, “Gabor-Zernike Features based Face Recognition Scheme,” *Int. J. Imaging Robot*, vol. 16, pp. 118–131, 2015.
- [28] C. Ding, J. Choi, D. Tao, and L. S. Davis, “Multi-Directional Multi-Level Dual-Cross Patterns for Robust Face Recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 3, pp. 518–531, 2016.
- [29] B. Amos, B. Ludwiczuk, and M. Satyanarayanan, “OpenFace: A general-purpose face recognition library with mobile applications.” CMU-CS-16-118, CMU School of Computer Science, 2016.
- [30] F. Schroff, D. Kalenichenko, and J. Philbin, “FaceNet: A unified embedding for face recognition and clustering,” *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 815–823, 2015.
- [31] A. Communications, “Facial Recognition in video surveillance,” <https://www.axis.com/es-mx/solutions-by-application/facial-recognition>, 2017.
- [32] Herta, “BioSurveillance,” <http://www.hertasecurity.com/productos/biosurveillance>, 2017.
- [33] OpenCV, “Histogram Equalization,” 2018. [Online]. Available: https://docs.opencv.org/3.1.0/d5/daf/tutorial_py_histogram_equalization.html.
- [34] N. S. Vu and A. Caplier, “Illumination-robust face recognition using retina modeling,” in *2009 16th IEEE International Conference on Image Processing (ICIP)*, pp. 3289–3292, 2009.
- [35] J. Héroult and B. Durette, “Modeling Visual Perception for Image Processing,” *Computational and Ambient Intelligence*, pp. 662–675, 2007.
- [36] D. E. King, “Dlibml: A Machine Learning Toolkit,” *Journal of Machine Learning Research*, vol.

- 10, pp. 1755–1758, 2009.
- [37] T. Edmunds and A. Caplier, “Motion-based countermeasure against photo and video spoofing attacks in face recognition,” *Journal of Visual Communication and Image Representation*, vol. 50, pp. 314–332, 2018.
- [38] Zhou Wang and I. Alan C. Bovik, Fellow, “Image Quality Assessment: From Error Visibility to Structural Similarity,” *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 1–14, 2004.
- [39] D. Basso, “Propuesta de Métricas para Proyectos de Explotación de Información,” in *Revista Latinoamericana de Ingeniería de Software*, 2015, vol. 2, no. 4, p. 157.
- [40] V. Kazemi and J. Sullivan, “One Millisecond Face Alignment with an Ensemble of Regression Trees,” *In Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1867–1874, 2014.
- [41] Intel, “OpenCV,” 2017. [Online]. Available: <https://www.opencv.org/>.
- [42] D. King, “dlib C++ Library,” 2015. [Online]. Available: [Www.Dlib.Net](http://www.dlib.net).
- [43] W. Valderrama, A. Magadán, and M. Mejía, “Comparación de Algoritmos de Clasificación con Reducción de Variables para Reconocimiento Facial,” in *Herramientas y tecnologías específicas para beneficio de las instituciones y comunidades, IEEE*, 2018, no. 978-607-95255-8-3, pp. 191–195.
- [44] Microsoft Corp., “Visual Studio,” *Visualstudio.Com*, 2015. [Online]. Available: <https://www.visualstudio.com/>.
- [45] Q. D. Frameworks, “Qt,” 2018. [Online]. Available: <https://www.qt.io>.
- [46] D. C. E. Thomaz, “FEI Face Database,” *Image Processing Laboratory Department of Electrical Engineering Centro Universitario da FEI, São Bernardo do Campo, São Paulo, Brazil*, 2012. [Online]. Available: <https://fei.edu.br/~cet/facedatabase.html>.
- [47] S. Libor, “Collection of facial images: Faces94,” *Computer Vision Science and Research Projects, University of Essex, United Kingdom*, 2007. [Online]. Available: <http://cswww.essex.ac.uk/mv/allfaces/faces94>.
- [48] K. MARTIN and B. HOFFMAN, “Mastering CMake,” 2010. [Online]. Available: <https://cmake.org/>.