INSTITUTO POLITÉCNICO NACIONAL

UNIDAD PROFESIONAL INTERDISCIPLINARIA EN INGENIERÍA Y TECNOLOGÍAS AVANZADAS

Trabajo Terminal II

"Sistema de control de acceso de personas basado en reconocimiento facial"

Que para obtener el título de

"Ingeniero en Mecatrónica"

Presenta:

Daniel Miranda Badillo

Jorge Alejandro Peña Muñoz

Eduardo Cruz González

Asesores:

M. en C. Adrián Morales Blas

Dr. Erik Zamora Gómez

Mayo 2018

INSTITUTO POLITÉCNICO NACIONAL

UNIDAD PROFESIONAL INTERDISCIPLINARIA EN INGENIERÍA Y TECNOLOGÍAS AVANZADAS

Trabajo Terminal II

"Sistema de control de acceso de personas basado en reconocimiento facial"

Que para obtener el título de

"Ingeniero en Mecatrónica"

Presenta:

Daniel Miranda Badillo

Jorge Alejandro Peña Muñoz

Eduardo Cruz González

Asesores:

M. en C. Adrián Morales Blas

Presidente del Jurado

Dr. Erik Zamora Gómez

Profesor titular

Dra. Blanca Tovar Corona

Dr. Juan Luis Mata Machuca

Contenido

1. RESUMEN 6
2. OBJETIVOS 7
2.1 Objetivo general 7
2.2 Objetivos específicos 7
3. INTRODUCCIÓN 8
4. JUSTIFICACIÓN 8
5. ESTADO DEL ARTE 9
5.1 Antecedentes internacionales 10
5.2 Antecedentes nacionales 11
6. MARCO TEÓRICO 12
6.1 Enfoques en el reconocimiento facial según la técnica de análisis empleada 12
6.2 Adquisición de imágenes 14
6.2.1 Sistemas de iluminación 15
6.2.2 El fondo 16
6.3 Detección de rostros genuinos 16
6.4 Clasificadores en cascada Haar (detección de rostros). 17
6.5 Técnicas de mejoramiento de imagen 19
6.5.1 Ecualización de histograma. 19
6.5.1 Ecualización de histograma adaptativo por límite de contraste: CLAHE (*"Contrast Limited Adaptive Histogram Equalization"*) 20
6.6 Extracción de características. 21
6.6.1 Análisis de componentes principales (PCA). 21
6.6.2 LDA- *Linear Discriminant Analysis "FIsherfaces"* 23
6.7 Métodos de clasificación. 24
6.7.1 SVM: Support Vector Machine 24
7. PLANTEAMIENTO DEL PROBLEMA 25
7.1 Iluminación 25
7.2 Posición y expresiones 26
7.3 Detección de rostro genuino 26
8. METODOLOGÍA 26
9. DESARROLLO 28
9.1 Selección de Microcomputadora. 28

9.2 Selección y puesta a punto de cámara. 29
9.3 Elementos adicionales para el control de acceso 31
9.6 Establecimiento de condiciones de entorno 32
9.6.1 Selección de iluminación 32
9.6.2 Normalización de iluminación 34
9.7 Diseño propuesto del sistema de control de acceso 34
9.7.1 Construcción de la cabina 35
9.7.2 Construcción de la carcasa del sistema. 36
9.8 Área de interés. 37
9.8.1 Detección de rostro dentro del área de interés. 37
9.8.2 Obtención de área de interés. 38
9.9 Normalización de imagen 39
9.3 Técnicas de detección de rostros genuinos 41
9.4 Selección del método a emplear para la detección de rostro genuino 42
9.4.1 Detección de rostro genuino mediante detección de parpadeo 42
9.4.2 Detección de rostro genuino mediante sensor de distancia 44
9.5 Esquema del sistema en ROS 45
9.10 Extracción de características 46
9.10.1 Extracción de características por análisis de componentes principales PCA o "Eigenfaces" 47
9.10.2 Extracción de características por Análisis lineal discriminante (LDA) o "Fisherfaces" 49
9.11 Clasificación 50
9.11.1 Clasificación con PCA como método de extracción de características. 51
9.11.2 Clasificación con LDA como método de extracción de características. 53
9.12 Desarrollo de interfaz gráfica de usuario (GUI) 55
10. ANÁLISIS DE RESULTADOS 59
10.1 Resultados PCA con 10 usuarios registrados. 59
10.2 Resultados LDA con 10 usuarios registrados. 60
10.3 Resultados PCA con 17 usuarios registrados. 62
11. GUÍA DE USO DE LA INTERFAZ DE USUARIO 63
12. CONCLUSIONES 67
13. TRABAJOS FUTUROS 68
14. REFERENCIAS 69

1. RESUMEN

En el presente trabajo se muestra el desarrollo del proyecto: “Sistema de control de acceso de personas basado en reconocimiento facial”, el cual consiste en la implementación de técnicas de visión artificial orientadas al reconocimiento facial, empleando una computadora de multiproceso heterogénea dentro de una cabina con iluminación normalizada, en donde se ha de llevar a cabo el proceso de reconocimiento facial. El proceso consiste en la adquisición y preprocesamiento de imágenes, la implementación de una técnica de extracción de características y un proceso de clasificación que determinará un resultado positivo o negativo para el control de acceso, en este caso una puerta.

Las etapas que involucran software en el sistema han sido desarrolladas e implementadas en el sistema operativo *ROS (Robotic Operating System)*, el cual está basado en una arquitectura de grafos donde el procesamiento toma lugar en nodos, en los cuales se han aplicado técnicas de preprocesamiento, tal como el algoritmo de ecualización adaptativa para el control de iluminación, que contrarresta pequeñas variaciones de iluminación en el rostro y genera las imágenes que han de ser empleadas en la extracción de características, para su posterior clasificación.

Asimismo, los algoritmos relacionados con el preprocesamiento, la aplicación de métodos de extracción de características y el mecanismo de clasificación, han sido desarrollados en lenguaje *Python*. Así como el manejo de información de la base de datos y el desarrollo de una interfaz gráfica que permitirá la manipulación del sistema desde un ambiente agradable y fácil de usar.

El trabajo realizado ha sido desarrollado en la plataforma de desarrollo colaborativo GitHub para el control de versiones, el cual está disponible en el siguiente enlace:

https://github.com/alexpm94/tt2_pack

2. OBJETIVOS

2.1 Objetivo general

Diseñar y construir un sistema de control de acceso personal mediante reconocimiento facial, implementando dos técnicas de extracción de características, en una tarjeta multiproceso y en un entorno semicontrolado.

2.2 Objetivos específicos

- Implementar un sistema de adquisición de imágenes que permita validar el acceso a usuarios registrados y capturar imágenes para definir un nuevo usuario.
- Diseñar e Implementar un entorno físico donde se realice la captura de imágenes.
- Diseñar una técnica de control de luminosidad que le permita al sistema adaptarse a pequeñas variaciones de iluminación y normalizar un rango de iluminación.
- Seleccionar y adaptar una tarjeta multiproceso que permita el manejo y acondicionamiento de datos y señales, además de garantizar la compatibilidad entre los subsistemas.
- Evaluar el funcionamiento global del sistema con el uso de 2 técnicas para la extracción de características que permitan la identificación genuina de rostro, con el fin de garantizar una eficiencia de al menos 70%. Se implementará solo la técnica que presente el mayor índice de identificación correcta.
- Implementar una base de datos que permita al sistema almacenar y acceder a la información asociada de al menos diez rostros.
- Diseñar y construir una interfaz visual que guíe al usuario en el proceso de identificación, desde el reconocimiento hasta la validación.
- Acoplar el sistema de reconocimiento facial a una puerta que permita el acceso en caso de validación, así como permitir la salida del personal sin necesidad de autenticación.

3. INTRODUCCIÓN

Como respuesta a la necesidad de seguridad y gestión de acceso de personas, se han diseñado diversos sistemas de control de acceso, algunas soluciones integran dispositivos electrónicos, software y dispositivos electromecánicos. Algunos de estos sistemas además de realizar la función de autenticación incorporan mejoras en la disminución de vulnerabilidades mediante el uso de tecnologías biométricas.

El reconocimiento facial destaca por ser una técnica con alta capacidad de respuesta ante otros tipos de identificación biométrica, dado que tiene una buena aceptación y suele ser útil para aplicaciones no intrusivas. [1] Por otro lado, las aplicaciones basadas en reconocimiento facial, presentan limitantes como la iluminación, envejecimiento, orientación de rostro y expresiones faciales. El trabajo que se presenta, pretende mostrar el funcionamiento de un sistema capaz de identificar la identidad de una persona mediante visión artificial. Se decidió realizar este proyecto con el objetivo de crear un dispositivo funcional capaz de realizar por completo el proceso de control de acceso y no solo la etapa de reconocimiento facial.

4. JUSTIFICACIÓN

Hoy en día existen diversos sistemas que pueden llevar a cabo la tarea de control de acceso de personal empleando tecnologías biométricas, este proyecto se centra en la aplicación de técnicas de visión artificial y tiene como finalidad el emplear métodos de reconocimiento facial basados en apariencias (métodos holísticos), donde se analiza la información del rostro completo como un conjunto y no se enfocan en aspectos tales como la distancia entre ojos, que corresponden a métodos basados en modelos (métodos analíticos) como se ha realizado en algunos proyectos anteriores.

El sistema propuesto involucra la integración de diversos elementos como son: el diseño de la estructura para montar los componentes para el reconocimiento facial, la iluminación y la conexión de una microcomputadora con distintos periféricos tales como son: cámara digital, pantalla táctil, tarjeta de programación para adquisición de datos y control de relés para el accionamiento del mecanismo de acceso. En cuanto a software involucra técnicas para el preprocesamiento, tratamiento de imágenes, extracción de características, mecanismo de clasificación, manejo de datos en formato CSV, utilización de librerías de código abierto como “scikit-learn” para el manejo de herramientas de machine Learning y el desarrollo de una interfaz gráfica, todo esto integrado en ROS que provee de herramientas que permiten la visualización, comunicación, administración de paquetes, entre otros.

A diferencia de los proyectos de reconocimiento desarrollados en la UPIITA, este sistema ha sido implementado en una tarjeta de multiproceso heterogénea en el sistema operativo *ROS* que presenta ventajas tales como:

- Permite cierto grado de autonomía entre procesos, es decir que se pueden realizar algoritmos en paralelo

- Se pueden realizar pruebas de algoritmos o técnicas de procesamiento de imágenes en cualquier momento sin necesidad de estar en el área de detección.
- Se pueden actualizar las bases de datos periódicamente con relativa facilidad a partir de la captura secuencial de imágenes.

Al desarrollar este proyecto se pretende construir un sistema que pueda excluir vulnerabilidades que presentan sistemas de reconocimiento facial existentes [2]. Hay sistemas que no garantizan la detección genuina de una persona, es decir es vulnerable ante una fotografía de alguna persona registrada, validando el acceso a una persona ficticia. Existen otros sistemas que no utilizan métodos de reconocimiento de patrones adecuados para el reconocimiento facial, por ejemplo, el método "*Adaboost*", éste realiza la clasificación de patrones de forma lineal.

5. ESTADO DEL ARTE

La tecnología biométrica es el desarrollo de aplicaciones (sistemas biométricos) que permiten llevar a cabo de manera automatizada la identificación y verificación de la identidad de los individuos a partir de rasgos únicos que estos poseen [3]. A continuación, se presentan de manera general las tecnologías biométricas con más presencia en el mercado.

- Reconocimiento de Huella Dactilar
- Reconocimiento de Iris y Retina
- Reconocimiento de la Geometría de la mano
- Reconocimiento de Firma escrita
- Reconocimiento de Voz
- Reconocimiento Facial

La visión artificial es una disciplina que forma parte del campo de la inteligencia artificial, haciendo uso de otras disciplinas como por ejemplo el aprendizaje automático y redes neuronales se han desarrollado grandes avances en tecnología habiéndose extendido sus aplicaciones, una de ellas es el reconocimiento facial donde se trabaja para mejorar el rendimiento de estos sistemas. Algunos sistemas comerciales internacionales existentes se presentan a continuación.

5.1 Antecedentes internacionales

FaceSentinel

Es un sistema biométrico desarrollado por la empresa británica "Aurora". **Figura 1.** Es un sistema que permite el acceso a ciertos usuarios para entrar a un lugar restringido. FaceSentinel fue creado para controlar el acceso de los usuarios de ciertas aerolíneas en aeropuertos. El motor de reconocimiento dentro del dispositivo usa métodos de aprendizaje basados *Deep Learning* [2].

FIGURA 1 SISTEMA FACESENTINEL

La empresa Aurora asegura que el tiempo de verificación es menor a 1 segundo. Se garantiza el acceso del usuario, aún si este se encuentra en una posición distinta, es decir, que no esté completamente de frente.

FaceSentinel también almacena imágenes en todos los intentos de acceso, de tal forma que no importa si la persona accede o no, el dispositivo guardará la imagen.

IFace 302

La empresa argentina "*Fingertech*" ofrece en su sitio web un modelo llamado *Iface* 302, y lo presenta como un equipo autónomo para el control de tiempos y asistencias mediante tecnología de huella dactilar y reconocimiento facial [4].**Figura 2.** Tiene una capacidad de identificar 400 rostros faciales y 5000 huellas dactilares y se ofrece en el mercado en un valor aproximado de $300 USD. Es un sistema que mediante algoritmos de visión es capaz de reconocer rostros, obtener y guardar datos en una base de datos. Si bien no es un sistema de control de acceso, tiene tecnología de características similares relacionadas a la identificación de rostros.

FIGURA 2 CONTROL DE ASISTENCIAS IFACE 302

Face ID F910

Otro dispositivo comercial que utiliza el reconocimiento facial como principal fuente biométrica para diferentes usos de control de personal es el "F910" (**Figura 3**) de la empresa *Face ID* [5] [6] que utiliza identificación biométrica facial 2D para reconocer al usuario de manera segura y fiable sin posibilidad de suplantación de identidad y además sin que exista contacto entre el usuario y el terminal, con esto la seguridad biométrica es perfectamente compatible con los requisitos de higiene.

FIGURA 3 FACE ID F910

5.2 Antecedentes nacionales

Reconocimiento de rostros utilizando análisis de componente principales: limitaciones del algoritmo.

Como expone Quezada en su tesis [7] su principal objetivo es comparar diferentes algoritmos de reconocimiento facial comerciales y concluir su eficiencia en MATLAB. En el trabajo de investigación se presentan diversos algoritmos basados en métodos estadísticos cómo redes neuronales y regresión, también se analizan algoritmos basados en "componentes principales". Finalmente, el autor postula un método que se basa en funciones polinomiales multivariadas de aproximación.

Destacamos que esta es una tesis de investigación. El autor desarrolla un nuevo método de reconocimiento facial pero no indica una aplicación específica en la que utilizará el método desarrollado.

Diseño e implementación de una herramienta de detección facial.

Esta tesis fue desarrollada por María Esther García Chang para obtener el grado de maestría en tecnología de cómputo, en el Centro de Innovación y Desarrollo Tecnológico en Cómputo (CIDETEC-IPN).

La principal aportación de este trabajo es identificar un método de detección de rostro en tiempo real, de tal forma que es posible detectar rostros en un video. El método elegido para detectar rostros en una imagen es el método "*Adaboost*", utilizando las librerías de código libre "OpenCV". En este trabajo únicamente se detectan rostros en una imagen, no se detectan rostros específicos.

Sistema de reconocimiento facial

En la UPIITA se pueden encontrar dos trabajos como antecedentes, en el primero de ellos se desarrolló un sistema que analiza los rostros de perfiles (mirando de frente, hacia arriba, hacia abajo, a la derecha y a la izquierda), realizando un análisis para cada uno, teniendo como objetivo aumentar la eficiencia en el reconocimiento [8]. Este sistema desarrolla el proceso de identificación y autenticación de rostro y establece como posible mejora su implementación ya que el control de acceso solo fue simulado. En este proyecto se emplearon redes neuronales de retropropagación y un clasificador euclidiano para la extracción de los vectores característicos de cada persona. Así para solucionar el problema de reconocimiento de rostro se combinaron invariantes de Hu y Yan obteniéndose una eficiencia del 82%.

Sistema de seguridad mediante el reconocimiento ocular

El segundo trata sobre un sistema de seguridad basado en el reconocimiento biométrico concentrándose únicamente en la zona donde se encuentran los ojos para el control del acceso de personas a áreas restringidas. [9] Consiste en la detección de puntos extraídos del área de la cara donde se encuentran los ojos

Este proyecto fue desarrollado en MATLAB, se implementó una base estructural para que el usuario coloque la cabeza para la captura de imagen, se integró el sistema en una base de datos desarrollada en MySQL y se implementó una interfaz gráfica de usuario desarrollada con la herramienta *GUI* de Matlab.

6. MARCO TEÓRICO

6.1 Enfoques en el reconocimiento facial según la técnica de análisis empleada.

Los sistemas que emplean reconocimiento facial se clasifican según la manera en la que los algoritmos se encargan de realizar el análisis de datos: métodos basados en características y métodos basados en apariencia. [10]

Los métodos basados en características, realizan el análisis de las propiedades y la geometría del rostro, tales como áreas, distancias y ángulos entre los elementos del rostro.

Los métodos basados en apariencia, conocidos también como métodos holísticos, consideran al rostro como un todo y realizan un análisis global utilizando herramientas estadísticas. Estos métodos buscan un nuevo subespacio de menor dimensión para proyectar los rostros. [11] Entre estos métodos se encuentran: el análisis de componentes principales (PCA), el análisis discriminante lineal (LDA), el análisis de componente independiente (ICA), el análisis de características locales (LFA), Tensorfaces, Redes neuronales (NN) y las máquinas de soporte (SVM).

Este proyecto está orientado a los métodos basados en apariencia, teniendo esto en cuenta se tiene que un sistema de reconocimiento facial en general sigue el esquema de la **Figura 4.**

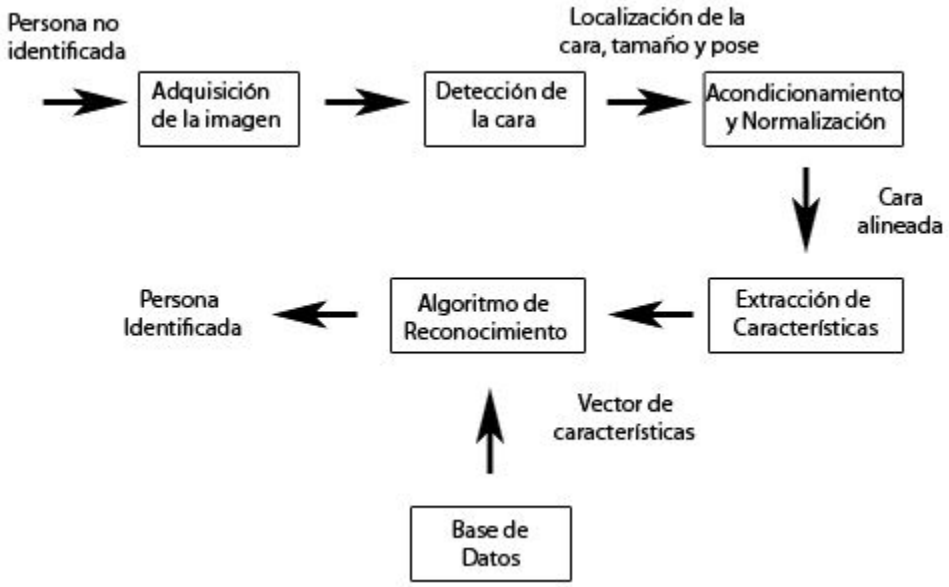


FIGURA 4 ELEMENTOS DE UN SISTEMA DE RECONOCIMIENTO FACIAL

Si se analiza cada etapa del proceso se pueden identificar varias dificultades para alcanzar un nivel de eficiencia óptimo.

La primera etapa dentro de un proceso de visión computacional es la etapa de adquisición que se refiere a la obtención de una imagen a partir de una cámara. Una correcta adquisición de la imagen supone un paso muy importante para que el proceso de

reconocimiento tenga éxito. Dentro de esta etapa existen múltiples factores que influyen directamente al proceso de captura de la imagen: el sistema hardware de visión artificial (cámara, tarjeta de adquisición de imagen, computadora y software) y el entorno y posicionamiento de los elementos (la iluminación, el fondo, posición correcta de la cámara, etc.)

Para la etapa de detección del rostro, se identifica si hay algún rostro en la imagen, en algunos proyectos se suele emplear un método de ventana deslizante, que consiste en hacer un barrido en la imagen con cuadrados de diferentes tamaños, si dentro del cuadrado se encuentra un rostro, éste se guarda y al finalizar el barrido se compara cuál fue el cuadrado que tiene el mayor porcentaje de detección. Para detectar que dentro del cuadrado haya un rostro, se usan algoritmos de clasificación utilizando métodos de "*Machine learning*" cómo SVM ("*Support Vector Machine"*) y redes neuronales [12].

La etapa de acondicionamiento y normalización es tal vez la más importante en la visión artificial, pues si no se tiene una imagen en donde la región de interés esté bien definida, es complicado obtener un buen resultado en el proceso de extracción de características. En esta etapa se debe realizar un tratamiento de las imágenes para compensar las posibles variaciones de iluminación, eliminar el ruido mediante filtros y mejorar el contraste.

La extracción de características ayuda a reducir el costo computacional, ya que sólo se procesa la región de interés. Para definir la región de interés se busca que las regiones sean homogéneas respecto a alguna característica, en este caso la región de interés es el rostro en la imagen. Se recorta la imagen para posteriormente procesar únicamente el rostro. Una vez que se definió la región de interés se procede segmentar la imagen, esto se refiere a subdividir la imagen.

Para la etapa extracción de características existen diversas soluciones que se diferencian por la robustez de los algoritmos. Entre los diferentes métodos se encuentran los siguientes [13]:

- Análisis de componentes principales
- Análisis de componente independientes
- Modelo de apariencia activa
- Análisis de discriminantes lineales

En la etapa de reconocimiento se suelen usar los siguientes métodos de clasificación [13]:

- SVM (*Support Vector Machine*)
- Redes Neuronales
- HMM (*Hidden Markov Models*)
- Algoritmos genéricos

En las redes neuronales convolucionales después de la etapa extracción de características, los datos llegan a la fase de clasificación. Una vez que las características extraídas son clasificadas se activa la neurona que corresponde al objeto deseado, en este caso el rostro. [14]

6.2 Adquisición de imágenes

La cámara: Su elección depende de múltiples factores, tales como el tipo de sensor que emplea, velocidad de captura, resolución, conexión, tamaño, costo, etc.

Tarjeta de adquisición de imagen: Otro aspecto fundamental a considerar, dentro del sistema de visión artificial, es el de la tarjeta de captura de imagen. Existen una gran variedad de tarjetas en el mercado orientadas a aplicaciones profesionales, con gran cantidad de características que hace que sea difícil seleccionar unas u otras. Algunas minicomputadoras de uso general están acondicionadas para realizar en la misma, la adquisición y procesamiento de imagen, simplificando el problema. Se deben considerar aspectos como la capacidad de cálculo y velocidad de procesamiento.

FIGURA 5 RASPBERRY PI. TARJETA DE DESARROLLO POPULAR USADA EN LA ADQUISICIÓN Y PROCESAMIENTO DE IMÁGENES [15]

Elección del software de programación: Tan importante como la elección del hardware es la del software a utilizar. Esto incluye desde el lenguaje de programación hasta el sistema operativo donde se desea compilar los códigos de los programas que se desarrollen. Se deben tener en cuenta factores como la compatibilidad, distribución de código abierto y recursos disponibles tales como librerías de desarrollo.

Entorno de adquisición: En el entorno de adquisición de imágenes existen variables que influyen en la calidad de la imagen tales como la iluminación, el fondo, la posiciónn de la cámara, entre otras. Las condiciones del entorno suponen una limitante constante en todos los sistemas de reconocimiento facial, por lo tanto, es importante establecer condiciones a las que se someterá el sistema de reconocimiento facial.

Iluminación: La iluminación se lleva a cabo mediante un conjunto de artefactos que producen radiación luminosa que incide sobre los objetos a visualizar. Se puede citar algunos elementos como lámparas, pantallas fotográficas, filtros de luz, etc. Aunque en fotografía y óptica se hablan de varias técnicas, existen fundamentalmente dos formas de iluminación.

Iluminación frontal: Donde la luz incide directamente sobre el objeto, ya sea verticalmente, horizontalmente, de forma oblicua o de forma difusa.

Iluminación trasera o retroiluminación: Donde se ilumina una pantalla de forma que lo que se resalta ES únicamente el contorno del objeto. Este objeto puede estar delante o detrás de la pantalla.

La iluminación frontal como se muestra en la **Figura 6**, permite distinguir los detalles de los objetos, así como su forma, permitiendo extraer más parámetros de cada objeto como su color, textura, detalles internos, etc. Por otro lado, la retroiluminación sólo sirve para simplificar la detección de contornos.

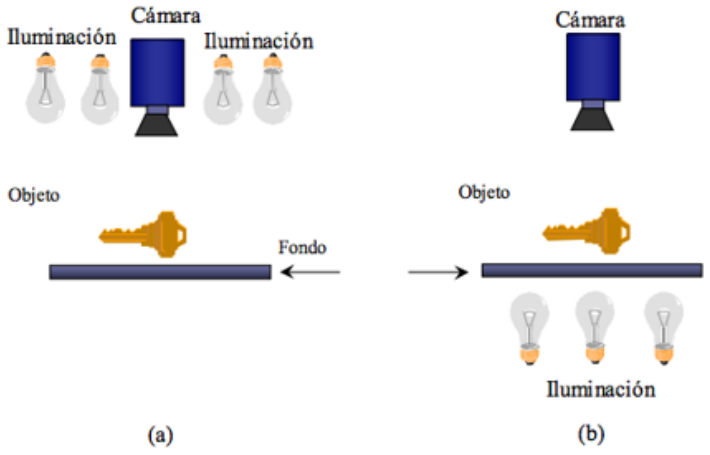


FIGURA 6 TIPOS DE ILUMINACIÓN. A) FRONTAL, B) RETROILUMINACIÓN [16]

6.2.1 Sistemas de iluminación

La selección de la iluminación más adecuada para cada aplicación depende del tamaño, las características de la superficie, y de las necesidades del sistema. A continuación, se presentan de manera general, las aplicaciones de cada sistema de iluminación.

Iluminador anular: Los iluminadores anulares o anillos de luz proporcionan una gran intensidad de luz uniforme, por lo que son adecuados para iluminar objetos de reducidas dimensiones. Se les puede acoplar filtros de colores, polarizadores, analizadores, y difusores para aumentar el efecto difusor.

FIGURA 7 ILUMINADOR ANULAR

Paneles difusos: Consisten en una placa iluminada internamente por *LED*, contiene una placa difusora que distribuye la luz de una manera uniforme, los dispositivos de mayor tamaño crean un mayor efecto difusor, que contrarresta las sombras y disminuye la concentración de iluminación. Suelen ser empleadas para iluminar grandes superficies.

FIGURA 8 ILUMINADORES DE TIPO PANEL

6.2.2 El fondo

El fondo de la escena cumple un papel importante cuando se trata de simplificar la segmentación del rostro. Utilizando iluminación frontal, el fondo debe de ser lo más opaco posible evitando todo reflejo. El color negro opaco suele ser el más utilizado además del verde oscuro. **Figura 9**.

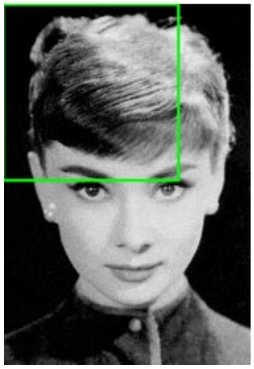


FIGURA 9 IMAGEN CON UN ROSTRO CON UN FONDO ADECUADO

6.3 Detección de rostros genuinos

A pesar del desarrollo en sistemas de reconocimiento facial, se presentan algunas vulnerabilidades donde los sistemas pueden reconocer que la imagen de un rostro corresponde en efecto a una persona, pero pasan por alto que se puede tratar de una fotografía o una imitación de un rostro. Este tipo de ataques a sistemas de reconocimiento facial mediante imágenes de rostros obtenidos a partir del robo de identidad es conocido como "*Face spoofing attack*" o "ataque de imitación" [17].

Con el fin de disminuir este tipo de vulnerabilidades se han propuesto técnicas que tienen como objetivo el identificar que la imagen que se presenta ante la cámara no es una fotografía, sino que se trata de una persona genuina. Las técnicas que se han empleado se pueden clasificar según la información de la imagen que se procesa como son:

Métodos basados en movimiento

En general este tipo de algoritmos intentan identificar que un rostro es genuino a partir del análisis del movimiento generado por la persona mediante secuencias de fotografías o videos. Por ejemplo, en [18] se tomó como referencia que el parpadeo en una persona ocurre aproximadamente entre cada 2 y 4 segundos, y se propuso un detector de parpadeos. También se han realizado métodos de análisis de movimiento en la boca o análisis de flujo óptico.

Comparado con otros métodos, los algoritmos basados en análisis de movimiento pueden obtener buenos resultados en la detección de imágenes estáticas en 2D, sin embargo, requieren de mayor tiempo para completar el análisis ya que están sujetos a la espera de detectar movimiento, por ejemplo, un parpadeo que puede tardar 3 segundos en promedio.

Métodos basados en análisis de textura

Este tipo de métodos se centra en las características de la piel en una imagen, tales como la textura y la reflectancia, esto suponiendo que existe diferencia entre las propiedades de un rostro real y una impresión. Por ejemplo, en [18] se realizó un método de detección de patrones de textura en imágenes a partir de defectos de impresión o imágenes borrosas, las cuales no corresponderían a un rostro genuino. Este método se basa en el análisis del espectro 2D de Fourier de imágenes de rostro genuino y de impresiones. El método funciona bien cuando existen marcadas diferencias o defectos de impresión en una imagen, pero no es fiable si se presenta una imagen de alta resolución. Comparado con otros métodos los algoritmos basados en el análisis de textura son rápidos en procesamiento, sin embargo, presentan grandes problemas ante variaciones de iluminación e imágenes de alta calidad.

Métodos basados en Hardware

Hasta la fecha no hay muchos sistemas basados en este tipo de métodos, se han propuesto algunos basados en imágenes fuera del espectro visual, tales como cámaras de profundidad, cámaras de infrarrojo complementario (CIR), en las cuales se compara la información de las caras reales y los elementos de ataques (fotografías, máscaras, pantallas), empleando sensores que funcionan en al menos dos longitudes de onda distintos. Por ejemplo, cámaras térmicas que comparan la temperatura facial y la temperatura irradiada por fotografías. También existen técnicas que combinan el análisis de movimiento de los labios con la voz de la persona. Este tipo de sistemas suelen proporcionar mejores resultados, pero requieren de piezas adicionales de hardware lo que se traduce en un incremento al costo del sistema.

6.4 Clasificadores en cascada Haar (detección de rostros).

Este clasificador se ha implementado en nuestro proyecto con la finalidad de encontrar en la imagen un rostro, fue propuesto por Viola y Jones en el cual, a partir de características locales de la imagen basadas en el cambio de intensidad, se podía desarrollar un detector de rostros muy robusto, usando filtros Haar (**Figura 10**) de diversos tamaños que recorren

la imagen, si dicho valor está por encima de un cierto umbral, esa zona de la imagen se clasificará como un rostro. [18]

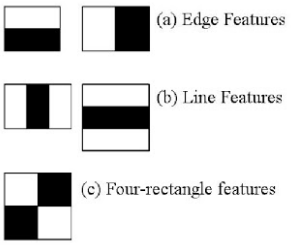


FIGURA 10 FILTROS HAAR [19]

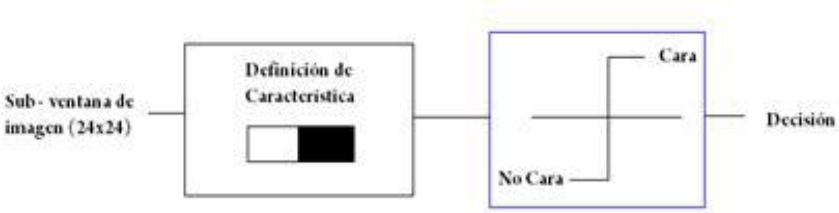


FIGURA 11 UMBRAL DECISIÓN HAAR [18]

Usar estos clasificadores individualmente da resultados pobres, la solución de Viola y Jones fue concatenar varios de estos clasificadores para obtener un clasificador robusto con una precisión del 99.9% de acierto, pero esto acarreaba un porcentaje elevado de detección de falsos positivos, más de 30%. [18]. Para solucionar esto los autores, proponen una concatenación de estos clasificadores fuertes, entrenando con todos los ejemplos que la etapa anterior no logro clasificar de manera correcta, teniendo un sistema similar al mostrado en la **Figura 12**.

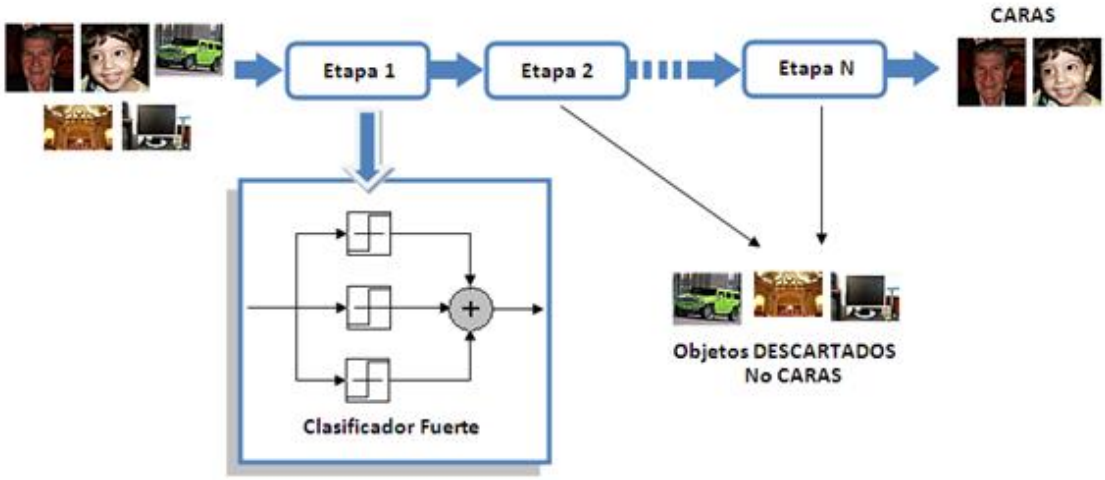


FIGURA 12 SISTEMA DE DETECCIÓN DE ROSTROS TIPO *CASCADE-HAAR* [18]

6.5 Técnicas de mejoramiento de imagen

6.5.1 Ecualización de histograma.

Un histograma es una representación gráfica de todos los pixeles de una imagen, donde se indica la frecuencia con la que se repiten los valores de brillo de los pixeles. Por ejemplo, en una imagen en escala de grises, el valor de los píxeles está en el rango de 0 a 255; si es una imagen con mucho brillo, significa que la mayoría de los píxeles están en un valor cercano a 255.

La ecualización de histograma se refiere a distribuir el histograma en el mayor rango de valores posibles como se muestra en la **Figura 13.**

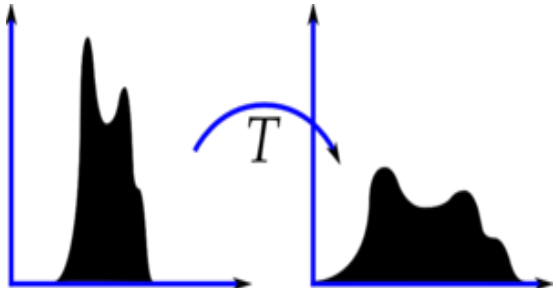


FIGURA 13 TRANSFORMACIÓN DE HISTOGRAMA [18]

FIGURA 14 IMAGEN ORIGINAL E IMAGEN CON ECUALIZACIÓN [19]

El procedimiento para realizar una ecualización es el siguiente [20]:

El histograma en una imagen está definido por:

Supongamos f, una imagen dada con rango de intensidad de 0 a $L-1$. L es el número de valores de intensidad posibles. La variable p denota el histograma normalizado de f.

$$p_n = \frac{numero\ de\ pixeles\ con\ intensidad\ n}{numero\ total\ de\ pixeles}$$

$n = 0,1,\dots,L-1$ Donde L suele estar en el rango $[0,255]$

El histograma ecualizado es:

$$g_{i,j} = floor((L-1)\sum_{n=0}^{f_{i,j}} p_n)$$

Donde $floor()$ redondea hacia abajo hacia el entero más cercano. Es equivalente a transformar la intensidad de los pixeles, k, por la función.

$$T(k) = floor((L-1)\sum_{n=0}^{k} p_n)$$

La ecualización se utiliza para mejorar el contraste de una imagen. Si una imagen es muy brillante, la ecualización tiende a oscurecerla y viceversa. Este método es eficiente cuando se aplica a una región particular de la imagen. Por ejemplo, si la imagen contiene tanto regiones oscuras como brillantes, lo que hace la ecualización es resaltar dichas diferencias.

6.5.1 Ecualización de histograma adaptativo por límite de contraste: CLAHE (*"Contrast Limited Adaptive Histogram Equalization"*)

Dentro del conjunto de imágenes que se encuentran en **Figura 15,** en la imagen original (**Figura 15 a**) se puede observar que el fondo es un poco oscuro; se puede pensar en aplicar una ecualización para mejorar el contraste.

En la **Figura 15 b** se observa la imagen después de aplicarle ecualización; sin embargo, se aprecia que la diferencia en brillo entre el maniquí y el fondo es aún más grande y se ha perdido información en el rostro del maniquí, por lo que no siempre es conveniente aplicar esta transformación. Aquí es donde entra la importancia del CLAHE.

Como se mencionó previamente, la ecualización es eficiente cuando se aplica a una región en particular, donde el valor de los píxeles no varíe demasiado para poder mejorar el contraste. El método CLAHE consiste en dividir la imagen en regiones y aplicar una ecualización individual a cada una de las regiones. En la **Figura 15 c** se puede observar el resultado de aplicar CLAHE, se notan más los detalles de los objetos y el fondo es más contrastante que en la imagen original.

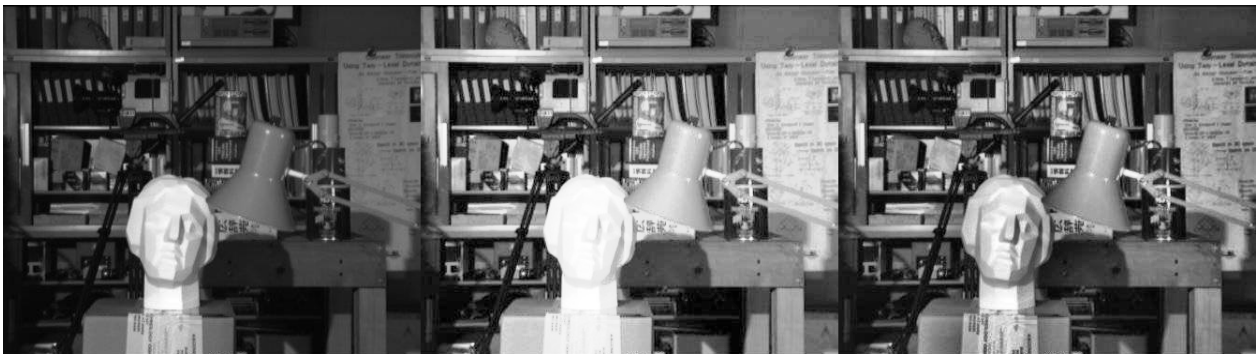


FIGURA 15 A) IMAGEN ORIGINAL B) IMAGEN DESPUÉS DE ECUALIZACIÓN C) IMAGEN DESPUÉS DE CLAHE [19]

Cabe destacar que para que el método CLAHE no amplifique el ruido que puede contener una imagen, se debe hacer un filtro para no realizar una ecualización en aquellos pixeles donde el contraste ya es elevado.

6.6 Extracción de características.

6.6.1 Análisis de componentes principales (PCA).

El análisis de componentes principales o PCA es un método clásico en el análisis de datos que tiene sus principales aplicaciones en la extracción de características y compresión de datos. Básicamente es un equivalente a la transformada de *Karhunen–Loeve* [21]. El PCA se presenta como una herramienta robusta en la identificación de patrones la cual permite comprimir datos o reducir la dimensión de una imagen conservando la información más relevante.

Sirovich [22] y Kirbi [21] fueron los primeros en usar PCA eficientemente para hacer representaciones de rostros humanos. Ellos sostenían que cualquier rostro podía ser reconstruido a partir de una imagen promedio y la suma ponderada de un pequeño conjunto de imágenes que forman una base de imágenes "Eigenimágenes". Basados en las publicaciones anteriores Turk y Pentland [23] presentaron en 1991 método conocido como "Eigenfaces." Desde entonces el PCA ha sido ampliamente investigado y se ha convertido en uno de los enfoques más exitosos en el reconocimiento facial [24].

Penev y Sirovich discutieron el problema de la dimensionalidad del espacio de rostros cuando se usa Eigenfaces para extracción de características. Zhao y Yang [25] intentaron explicar los efectos arbitrarios de la iluminación para sistemas de visión basados en PCA generando una forma de la matriz de covarianza para un caso particular de iluminación.

6.6.1.1 EIGENFACES – Descripción del proceso para el cálculo de PCA

1. Obtener un conjunto de M imágenes de rostro de tamaño NxN para cada una de las personas que se desee identificar

$$\Gamma_1, \Gamma_2, \dots \Gamma_M$$

2. Reorganizar cada imagen en un vector unidimensional de tamaño N^2 a partir de la concatenación de cada una de las filas de la imagen, para formar una matriz de tamaño $N^2 x M$ **Figura 16**.

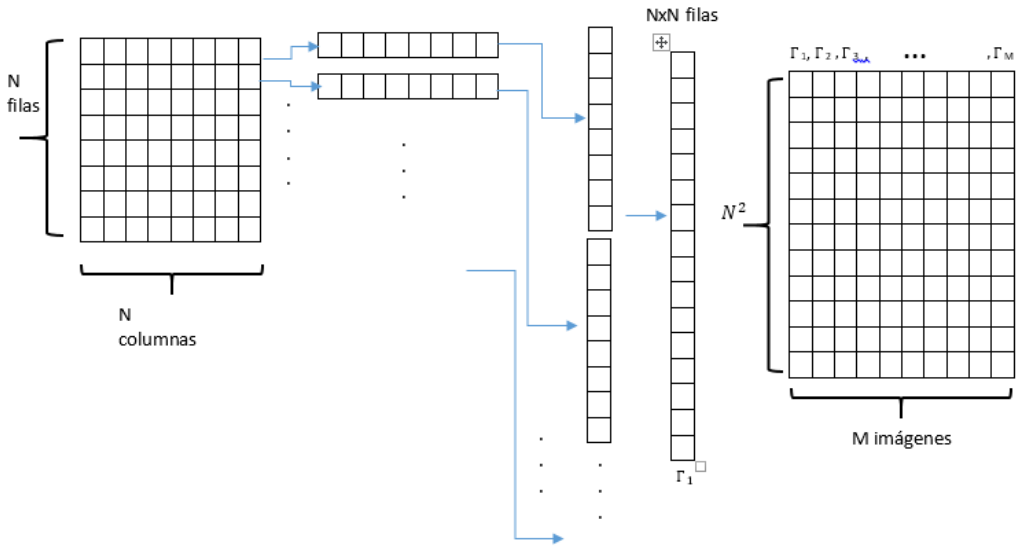


FIGURA 16 CONSTRUCCIÓN DE LA MATRIZ DE IMÁGENES CORRESPONDIENTE A UNA PERSONA M1

3. Obtener el rostro promedio a partir de la matriz anterior.

$$\mu = \frac{1}{M}\sum_{n=1}^{M}\Gamma_n$$

4. Restar el rostro promedio obtenido μ a cada una de las imágenes Γ_i para i= 1,2,…,M Obtener un nuevo conjunto de vectores Φ.

$$\Phi_i = \Gamma_i - \mu$$

5. Se obtiene la matriz A, de tamaño $\mathrm{N}^2\mathrm{xM}$

$$\Lambda = [\Phi_1 \quad \Phi_2 \quad \dots \quad \Phi_M]$$

6. Se obtiene la matriz de covarianza de la matrizΛ.

$$C = \frac{1}{M}\sum_{i=1}^{M}\Phi_i * {\Phi_i}^T = \Lambda\Lambda^T$$

7. Se calculan los eigenvectores v_i y eigenvalores λ_i de la matriz de covarianza C

$$Cv_i = \lambda_i v_i \qquad \text{i=1,2,…, N}$$

8. Ordenar los eigenvectores v_i de mayor a menor según su eigenvalor correspondiente λ_i; un eigenvalor alto significa que contiene la mayor información respecto a las imágenes del conjunto. Los componentes principales *k* son los eigenvectores correspondientes a los *k mayores* eigenvalores, solo se conservan los eigenvectores asociados a eigenvalores no-cero.

$$W = (v1, v2, \dots, v_k)$$

Como la matriz de covarianza "C" es de tamaño $N^2\ x\ M$ existe un problema en cuanto a dimensión, por ejemplo, si se tienen 400 imágenes de 100 x 100 pixeles, el cálculo de la matriz de covarianza $C = \Lambda\Lambda^T$ (con Λ de tamaño 10000 x 400) terminaría con una matriz de 10000 x 10000 elementos, aproximadamente 0.8 GB. [26] Este cálculo resultaría computacionalmente costoso de realizar. Existe una solución a este problema que fue expuesta por Turk y Pentland en [23] que señala que existe un teorema en algebra lineal que establece que los eigenvalores de $C = \Lambda\Lambda^T$ y los de C'= $\Lambda^T\Lambda$ son coincidentes. Por lo tanto, los eigenvectores de C son los mismos que los de D multiplicados por la matriz Λ.

De esta manera al emplear este teorema se puede crear el eigenespacio a partir de una matriz de tamaño $M\ x\ M$ que es significantemente de menor dimensión a la matriz de covarianza de tamaño $N^2\ x\ N^2$.

9. Se multiplica la transpuesta de la matriz W con la matriz Λ para formar la matriz del espacio transformado Ω_k, la cual contiene los datos de las imágenes proyectadas en el nuevo espacio de dimensión k.

$$\Omega_k = W^T\Lambda$$

Mediante la combinación lineal de los “eigenfaces” se puede reconstruir cualquiera de los M rostros, sin embargo, esta reconstrucción no es perfecta debido a que se limitó las dimensiones del espacio para mejorar la eficiencia computacional del sistema. [27] La reconstrucción de rostros de realizaría de la siguiente forma:

$$\Lambda = \mathrm{W}\Omega_k$$

Para obtener la imagen reconstruida

$$\Gamma_i = \Lambda_i + \mu$$

O bien

$$\Gamma_i = \mathrm{W}\Omega_k + \mu$$

Hoy en día hay investigadores trabajando en sistemas que puedan superar las restricciones de condiciones controladas que presentan los algoritmos de los sistemas actuales tales como cambios en la iluminación, rotación, traslación, expresiones faciales, pelo, gafas, maquillaje, etc. El diseño de algoritmos efectivos antes estos cambios están aún en fase de investigación. [28]

6.6.2 LDA- *Linear Discriminant Analysis “FIsherfaces”*

LDA fue inventado por el gran estadístico R.A. Fisher. Lo utilizo para clasificar flores en su artículo en 1936. LDA es una técnica de aprendizaje supervisado para clasificar datos. La idea central de LDA es obtener una proyección de los datos en un espacio de menor dimensión con el fin de que la separación de clases o personas sea más sencilla. Es una técnica supervisada lo que quiere decir que para obtener dicha proyección se debe entrenar al sistema con patrones etiquetados. A diferencia del PCA, el LDA no busca minimizar el error de la representación.

Se han hecho varias implementaciones de LDA, entre ellas se encuentra “Fisher LDA” [29], que en resumen busca maximizar la siguiente función objetivo:

$$J(\omega) = \frac{\omega^T S_B \omega}{\omega^T S_W \omega}$$

Donde $S_B$es la matriz de dispersión inter-clase y S_W es la matriz de dispersión intraclase.

$$S_B = \sum_c N_c(\mu_c - \mu)(\mu_c - \mu)^T$$

$$S_B = \sum_c \sum_{i \in c} N_c\,(x_i - \mu)(x_i - \mu)^T$$

Siendo μ_c la media de cada clase, μ la media de todos los datos, N_c la cantidad de patrones de la clase *c*. Fisher LDA busca encontrar el vector ω que maximiza la función objetivo que debe cumplir, esto se realiza mediante la obtención del promedio total:

$$\mu_c = \frac{1}{N_c} \sum_{i \in c} x_i$$

$$\bar{x} = \frac{1}{N}\sum_{i} x_i = \frac{1}{N}\sum_{c} N_c\, \mu_c$$

Se obtiene la matriz de proyección que permite pasar del problema de alta dimensionalidad a uno de baja. Cada una de las columnas que forman esta matriz serán las bases del nuevo subespacio, las cuales se conocen como "Fisherfaces". Estas bases son las que garantizan una máxima discriminación entre clases por encima de ser las que mejor describen los datos, a diferencia de PCA. [1]

Una de las limitaciones que tiene LDA es que requiere de matrices de dispersión no singulares. Esto en aplicaciones de reconocimiento facial es complicado debido a que en muchas ocasiones la alta dimensionalidad de las imágenes es mucho mayor que el número de imágenes lo que desemboca en un problema de matrices singulares. Para evitar este efecto, se utilizan diferentes aproximaciones como *pseudo-inverse* LDA [30], que utiliza la pseudoinversa de la matriz de covarianza en lugar de su inversa, y PCA + LDA [30], que realiza un reducción previa de los datos utilizando PCA.

6.7 Métodos de clasificación.

Este proceso tiene como objetivo verificar la compatibilidad de una secuencia de imágenes adquiridas con imágenes almacenadas previamente. Algunos de los métodos más empleados son los siguientes:

Clasificadores	Referencias
SVM – Support Vector Machines	[32] [33] [34]
Clasificadores binarios	[35]
Análisis discriminante lineal	[36] [37]
Adaboost	[38]

TABLA 1 CLASIFICADORES EMPLEADOS EN LA DETECCIÓN DE ROSTROS GENUINOS

6.7.1 SVM: Support Vector Machine

Los soportes vectoriales han sido utilizados para resolver problemas de reconocimiento de patrones. Este método de clasificación mapea los datos de entrada a un espacio dimensional mayor donde se pueda construir un hiperplano con la finalidad de separar las clases de manera lineal. A dicho mapeo se le conoce como *kernel.* La función *kernel* y sus parámetros son elegidos de tal forma que la frontera en la dimensión del vector de clasificación es minimizada. [39]

El método SVM funciona de la siguiente forma:

Dado un conjunto de datos de entrenamiento de N puntos $\{y_k, x_k\}_{k=1}^{N}$, donde $x_k \in R^n$ es el k ésimo patrón de entrada y $y_k \in R^n$ es el k ésimo patrón de salida. El método de vectores de soporte de un clasificador tiene la siguiente forma:

$$y(x) = sign\left[\sum_{k=1}^{N} \alpha_k y_k \Phi(x, x_k) + b\right]$$

Donde α_k es un real positivo y b es un real constante.

Uso de *kernel*

Una SVM sólo es capaz de clasificar problemas linealmente separables. Los datos que se observan en la **Figura 17 A**, no son linealmente separables. La línea en rojo corresponde a la frontera de decisión ideal, la cual maximiza la distancia entre ésta el conjunto de datos.

Como se mencionó, el método SVM sólo clasifica conjuntos linealmente separables, entonces la línea en amarillo de la **Figura 17 A,** corresponde al mejor resultado que podría dar una SVM, pero es notable que existe al menos un punto que está mal clasificado y además la distancia entre del conjunto de datos y la frontera de decisión no es la máxima. Aquí es dónde radica la importancia de utilizar una función *kernel*. Se realiza un mapeo a otro espacio vectorial, dónde los datos sean linealmente separables, como se muestra en la **Figura 17 B**. Una vez que se encontró la frontera de decisión que maximiza el margen, se regresa al espacio original, resultando como frontera de decisión la línea en rojo de la **Figura 17 A**.

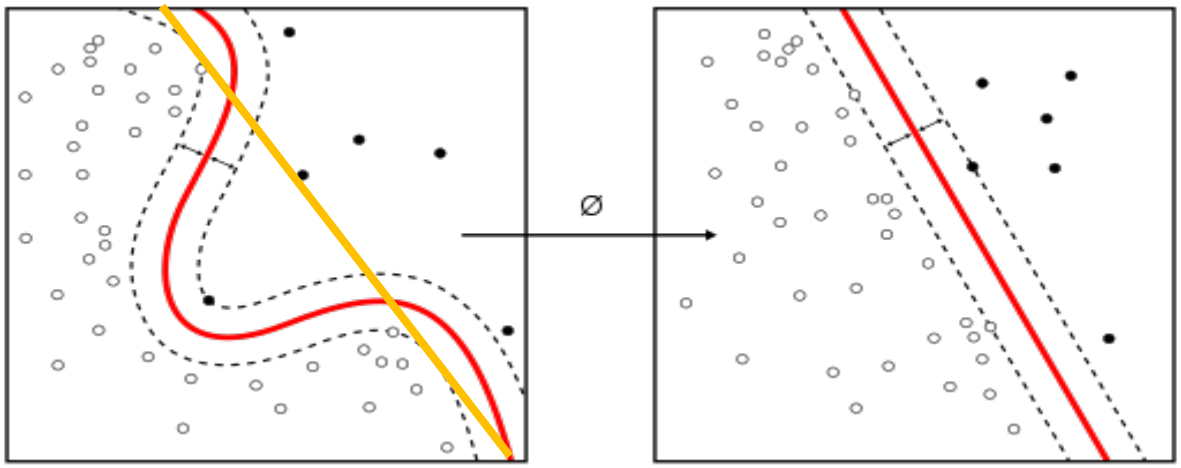


FIGURA 17 A) DATOS EN EL ESPACIO ORIGINAL [40] B) MAPEO DE DATOS USANDO KERNEL [40]

7. PLANTEAMIENTO DEL PROBLEMA

7.1 Iluminación

La iluminación en un sistema que implemente visión artificial es un factor importante ya que suele influir en la resolución de problemas, en ocasiones cuando ésta no se contempla se pueden acarrear problemas a las etapas de procesamiento y clasificación en un sistema de visión artificial. El propósito de la iluminación es el de controlar la forma en que la cámara adquiere información del rostro tratando que la imagen capturada cumpla con ciertas características tales como: disminución de sombras y luminosidad constante. El no acotar estas características puede derivar en dificultar el procesamiento de imágenes o disminuir la efectividad del sistema.

Se prevé que la variación en la iluminación tendrá que tomarse en cuenta como un factor importante en el desarrollo del proyecto, ya que, aunque no se puede tener siempre una iluminación idéntica, al emplearse condiciones controladas mediante una cabina e iluminación difusa se pretende hacer homogéneas las imágenes en la adquisición de datos.

7.2 Posición y expresiones

Uno de los problemas que suelen surgir en sistemas de reconocimiento facial tiene que ver con las diferencias en la orientación de la cabeza, así como algunas expresiones y complementos como son el portar lentes, o el crecimiento del vello facial, por ejemplo.

La variación de la iluminación es uno de los principales problemas. Se considera un sistema de iluminación semicontrolado, por lo que puede haber influencia de luz externa. El diseño de la cabina deja pasar luz sólo de un extremo, por lo que el rostro de la persona a identificar puede estar más iluminado de un lado que del otro, como se muestra en la **Figura 18**.

FIGURA 18 IMAGEN OBTENIDA DEL ENTORNO SIMULADO

Se puede observar que el rostro del individuo a detectar está más iluminado del extremo derecho. Uno de los métodos de extracción de características que se utilizará es PCA, el cuál es muy susceptible a variaciones de iluminación. Por esta razón las imágenes deben tener una iluminación homogénea en todo momento.

7.3 Detección de rostro genuino

En el caso de que se pretenda de engañar al sistema al presentar a la cámara una imagen 2D se trata de una vulnerabilidad conocida como como *spoofing attack* o ataque de imitación, el cual es común en los sistemas de reconocimiento facial. Uno de los problemas identificados es emplear un método o combinación de ellos ya sea análisis de movimiento, análisis de textura o empleando hardware adicional como un sensor para disminuir esta vulnerabilidad en particular.

8. METODOLOGÍA

Para la realización de este trabajo se ha planteado una organización como la mostrada en el esquema de la **Figura 19,** que permite dividir las tareas a realizar dentro de áreas funcionales, estas funciones son relativamente independientes del resto, y cada una de ellas está orientada al cumplimiento de los objetivos particulares. Se han identificado 7 áreas funcionales:

- Adquisición de imagen
- Detección de rostro
- Acondicionamiento y normalización de imagen
- Extracción de características

- Gestión de la base de datos
- Reconocimiento facial
- Control de acceso

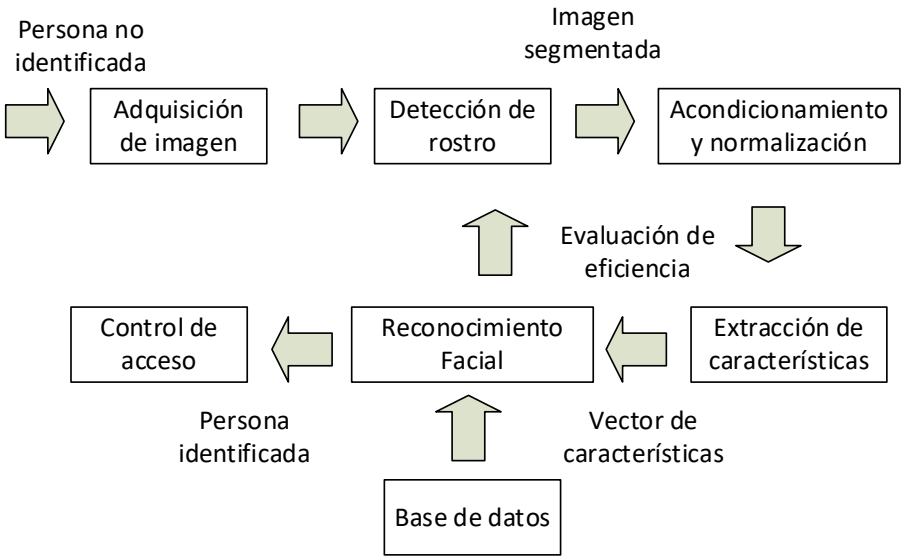


FIGURA 19 ORGANIZACIÓN FUNCIONAL DEL SISTEMA

Se ha señalado una etapa de evaluación ya que parte del proyecto es evaluar el funcionamiento con distintos métodos de extracción de características e implementar el que tenga mayor eficiencia para su aplicación final.

Las actividades más importantes desarrolladas en este proyecto según las áreas funcionales que lo conforman se ven representadas en la **Tabla 2**:

Área funcional	Tareas / Actividades
Adquisición de imágenes	• Selección de microcomputadora • Selección y puesta a punto de la cámara. • Conexión de periféricos y elementos de hardware • Establecimiento de condiciones del entorno de detección • Selección de iluminación • Diseño de cabina
Detección de rostro	• Detección del área de interés • Obtención del área de interés
Acondicionamiento y normalización	• Aplicación de método adaptativo de ecualización
Extracción de características	• Investigación sobre métodos de extracción de características. • Extracción de características por método de análisis de componentes principales “PCA”. • Extracción de características por método de análisis discriminante ideal “LDA”. • Implementación en Python de métodos de extracción de características.

	• Evaluación y comparación de resultados obtenidos a partir de la implementación de los métodos
Detección de rostro	• Investigación y selección de técnica para la solución a la vulnerabilidad de rostro genuino en 2D. • Implementación del método seleccionado • Evaluación del método empleado
Base de datos	• Investigación para seleccionar el tipo de base de datos a emplear • Diseño y normalización de la base de datos • Desarrollo e implementación de programas para manejo de base de datos • Ampliación de la base de datos
Control de acceso - Interfaz	• Investigación de interfaces de usuario compatibles con *ROS* • Elección de las herramientas para el diseño de la interfaz • Programación de la interfaz • Implementación e integración de la interfaz.

TABLA 2 DIVISIÓN DEL PROYECTO POR ÁREAS FUNCIONALES

9. DESARROLLO

9.1 Selección de Microcomputadora.

Existen muchas y variadas opciones de tarjetas de desarrollo que se podrían utilizar para el proyecto, pero solo una con procesador de multiproceso heterogéneo, que es la ODROID-XU4, se realizó una comparativa con la tarjeta de desarrollo de más uso en la actualidad que es la *Raspberry Pi 3 Model B*, esto con la finalidad de justificar la decisión de usar una tarjeta con procesador de multiproceso heterogéneo como lo es la ODROID-XU4.

Característica	1) Raspberry Pi 3 Model B [41]	2) Odroid XU4 [42]	Elección
Procesador	Quad Core 1.2GHz Broadcom BCM2837 64bit CPU [41]	Octa Core Samsung Exynos5422 4 x Cortex-A15 a 2GHz 4 x Cortex-A7 a 1.4GHz (HMP) [42]	2
Memoria RAM	1GB RAM [41]	2Gbyte LPDDR3 [42]	2
Tipo de almacenamiento	Micro SD [41]	Micro SD 64GB eMMC 64GB [42]	2
Puertos USB	4 USB 2.0 [41]	2 x USB3.0 Host 1 x USB2.0 Host [42]	2

Conectividad	WIFI Ethernet [41]	Ethernet LAN 10/100Mbps y Gigabit [42]	1
Soporte	Excelente	Buena	1

TABLA 3 COMPARATIVA RASPBERRY VS ODROID XU4

Para la elección se prestó atención en la velocidad de procesamiento, así como en la forma que se hace, recordemos que se realizará procesamiento de imágenes, extracción de características, toma de decisiones, todo esto es trabajo del procesador y se hará en menor tiempo, ya que no se quiere un sistema lento, esto se logró utilizando una memoria de alta velocidad eMMC en donde se instaló el sistema operativo y se almacenaron las imágenes de nuestra base de datos, además cuanta con un procesador HMP a una velocidad de 2 GHz.

La única ventaja significativa de la Raspberry sobre la ODROID es el soporte y la comunidad activa que se encuentra al tanto de los problemas que se tienen es por eso que muchas personas optan por lo más común. [43]

9.2 Selección y puesta a punto de cámara.

Se eligió que una cámara RGB es ideal para nuestro proyecto, pero en el mercado hay miles de opciones diferentes de cámaras RGB unas con más prestaciones que otras, aquí se enlistan algunas de las características se buscan en la cámara RGB que facilitan algunos pasos en el procesamiento digital de la imagen:

- Autoenfoque
- Ajuste de exposición automático
- Uso en diferentes sistemas operativos
- 30 FPS (*Frames* por segundo)

Son las características más importantes que se buscan en una cámara, comparando precios y calidades se eligió la cámara *Microsoft LifeCam Cinema*, que ofrece las características anteriores y algunas más, como lo son el grabar a 720p, el gran angular, y buena calidad de materiales.

FIGURA 20 ESPECIFICACIONES E IMAGEN DE CÁMARA *LIFECAM* DE MICROSOFT [44]

Para la puesta a punto de la cámara se trabajó en el funcionamiento de ésta en el entorno de ROS, para lo cual se descargó el paquete “usb_cam” que permite el uso de cámaras conectadas a la computadora mediante una interfaz USB, con esto se publican los datos que se están obteniendo de la cámara y se leen los datos con el visualizador de *ROS* conocido como “RViz” (**Figura 21**) que se suscribe a los datos publicados y se muestran.

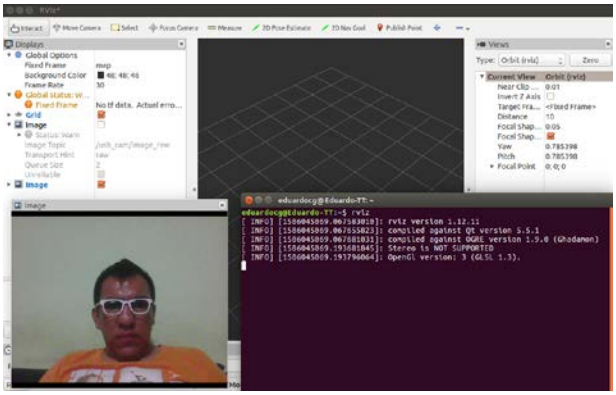


FIGURA 21 PUESTA A PUNTO CÁMARA.

Con lo anterior queda listo el sistema de adquisición de imágenes que nos permitirá hacer el reconocimiento facial además de agregar a nuevos usuarios al sistema.

9.3 Elementos adicionales para el control de acceso

Además de la tarjeta de multiproceso heterogéneo y la cámara USB se han de emplear otros elementos de hardware como son: teclado y ratón, pantalla táctil, conexión a internet por USB, y una placa de programación Arduino mediante la cual se lleva a cabo la adquisición de datos de un sensor ultrasónico y se envían señales lógicas a un módulo de relevadores para el accionamiento de la contrachapa y el panel de iluminación. En la **Figura 22** se muestra un esquema general de estos componentes.

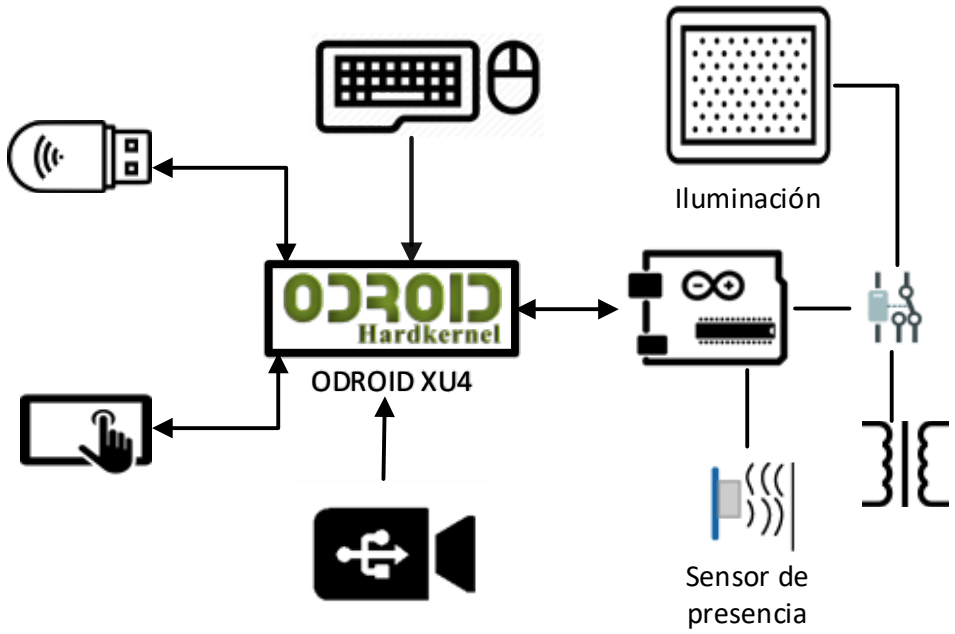


FIGURA 22 COMPONENTES DEL SISTEMA

El sensor empleado es el HC-SR04 que se emplea para detectar objetos y calcular la distancia a la que se encuentran, éste indicará si un objeto como un teléfono móvil o una fotografía es colocado a menos de 30cm de la cámara y no se dará el acceso. La distancia obtenida mediante el sensor ultrasónico también ha de servir para encender el panel de iluminación, esto si se detecta presencia de una persona frente a la cámara, en caso contrario se apagará después de 15 segundos después del último objeto reconocido.

Para el accionamiento de la contrachapa se empleó un transformador de 127 a 12V de C.A. conectado al solenoide que libera o bloquea el mecanismo de apertura de la contrachapa. Por otro lado, para el accionamiento del panel de iluminación se ha de conectar un balastro de 40 W.

La tarjeta de programación tiene comunicación con la Odroid - XU4 mediante la librería “roslib” de ROS, la cual permite utilizar la tarjeta Arduino como otro nodo de ROS que publica los datos del sensor y se suscribe a los tópicos de la Odroid que controlan el panel de iluminación y la contrachapa.

En la **Figura 23** se muestra el esquema de conexiones eléctricas que se realizó con los componentes mencionados.

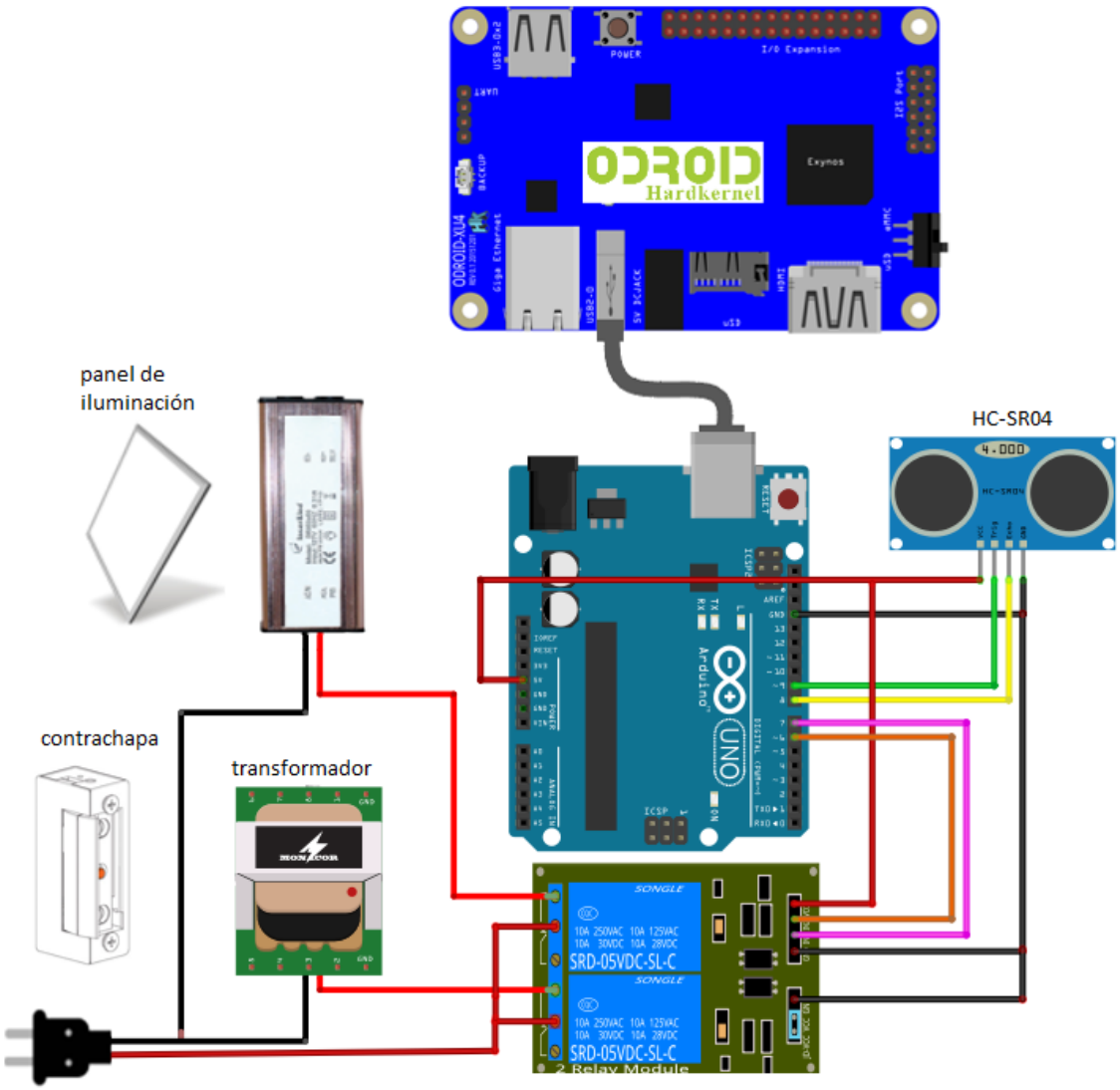


FIGURA 23 ESQUEMA DE CONEXIONES ELÉCTRICAS

9.6 Establecimiento de condiciones de entorno

9.6.1 Selección de iluminación

Los cambios de iluminación suelen hacer difícil el reconocimiento facial. [45] Por esto la iluminación en este proyecto es un factor importante, ya que se requiere disminuir en lo posible sombras y variaciones de luminosidad.

La retroiluminación suele emplearse en aplicaciones donde no es importante analizar la superficie del objeto que se desea analizar, por ejemplo, contornos o agujeros en una pieza; por lo que para este proyecto no es útil, siendo la iluminación frontal el tipo de iluminación adecuada para iluminar el rostro.

En cuanto a los dispositivos de iluminación frontal se encuentran: paneles y sistemas puntuales. En la **Tabla 4** se muestran los criterios a considerar para la elección del sistema que se empleará.

Criterio	Variable	Valor	Descripción
Disminución de sombras	Som	2.5	La iluminación disminuye las sombras presentes en el rostro.
Disminución de reflejos	Ref	2.5	La iluminación no produce reflejo o brillo que pueda dificultar su procesamiento.
Tamaño	Tam	2	Sus dimensiones son compatibles con la cabina propuesta, es decir, no ocupan más espacio extra.
Costo	Cos	1	Al ser un elemento auxiliar, se prefiere un dispositivo económico.
Iluminación excesiva	Lum	1	La luz emitida debe ser poco agresiva con el usuario.
Fácil instalación	Ins	1	Se prefiere un dispositivo que sea de fácil instalación.
	Total	10	

TABLA 4 CRITERIOS PARA SELECCIÓN DE ILUMINADOR

Tipo de iluminación	Criterios de evaluación						Total
	Som (2.5)	Ref (2.5)	Tam (2)	Cos (1)	Lum (1)	Ins (1.5)	
Iluminador anular	0.8	0.7	0.6	0.2	0.2	0.3	2.8
Iluminador lineal	0.3	0.2	0.2	0.2	0.2	0.35	1.45
Iluminación puntual	0.1	0.1	0.4	0.3	0.1	0.4	1.4
Iluminador indirecto	0.6	0.8	0.1	0.1	0.3	0.1	2.0
Panel de iluminación	0.7	0.7	0.7	0.2	0.2	0.35	2.85

TABLA 5 MATRIZ MORFOLÓGICA DE SELECCIÓN DE ILUMINACIÓN

El iluminador a emplear será de tipo panel instalando un difusor para disminuir la intensidad de la luz de forma que no sea molesta para el usuario. Este tipo de iluminador es fácil de instalar, ya existen distintos tamaños en versiones comerciales con LED.

FIGURA 24 PANEL DE ILUMINACIÓN COMERCIAL

9.6.2 Normalización de iluminación

El objetivo de la iluminación en este proyecto es resaltar los rasgos faciales que componen el área de interés, buscando obtener una iluminación constante tanto en la captura de imágenes para la base de datos como el reconocimiento.

La luz ideal corresponde con un tipo de iluminación difusa, en la que no existan sombras, brillos o reflejos en el objeto de interés. Se eligió una lámpara de tipo panel que tiene las siguientes características:

- Alimentación: 127 VCA
- Potencia: 40W
- Temperatura: 6500K
- IRC: 75

9.7 Diseño propuesto del sistema de control de acceso

Para el diseño del sistema se ha tomado en cuenta la selección de iluminación, cámara, pantalla, y características del entorno en que fue creada la base de datos.

La cabina tiene dos metros de altura, por un metro de ancho y un metro de profundidad, el cuadro de iluminación tiene 0.6 metros por lado, el sistema de reconocimiento está ubicado a una altura de 1.2 metros del suelo, la cámara y pantalla tienen una inclinación de 35° respecto a la vertical, esto con la finalidad de cubrir un rango de estaturas de personas entre 1.6 y 1.8 metros.

Para este prototipo se utilizó:

Madera para: la puerta, el marco de la puerta y el arco de la cabina.

Un cierrapuertas hidráulico, una cerradura eléctrica, un panel de iluminación de 40W.

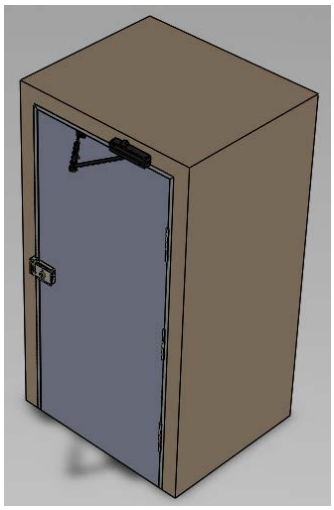


FIGURA 25 DISEÑO DE CABINA PARA CONTROL DE ACCESO

FIGURA 26 SIMULACIÓN DE USO DE LA CABINA

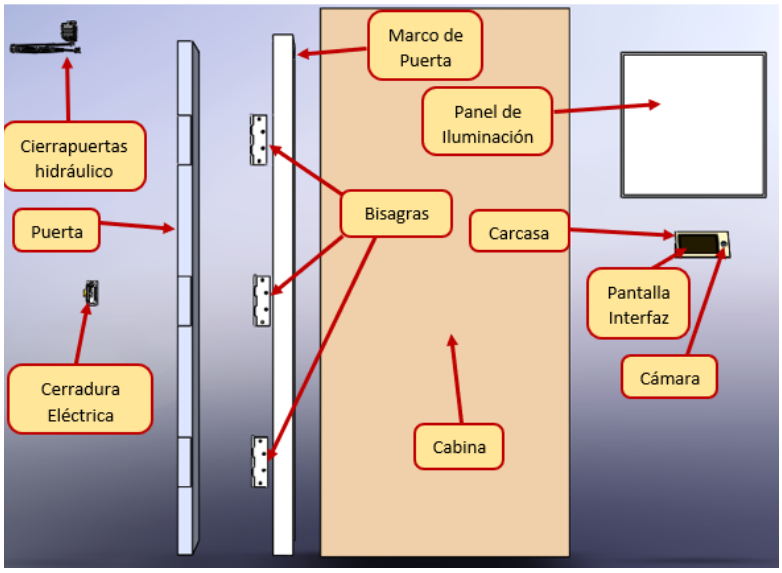


FIGURA 27 ELEMENTOS DE LA CABINA

9.7.1 Construcción de la cabina

Se decidió que el material de la cabina fuera madera, así que se mandó construir las partes de la cabina, siendo estas 5 partes principales, un piso, un techo, dos paredes, y una puerta, una vez armadas las partes, se realizó el diseño de la carcasa donde se encuentran los componentes principales del proyecto como son la cámara, la pantalla y la microcomputadora.

FIGURA 28 CABINA CONSTRUIDA, VISTA DESDE AFUERA.

FIGURA 29 VISTA DEL INTERIOR DE LA CABINA.

9.7.2 Construcción de la carcasa del sistema.

Para el diseño y construcción de la carcasa del sistema se tenían ciertos requerimientos, debería ser lo más compacta posible, además de que todos los componentes como pantalla, sensor de profundidad y cámara se pudieran acoplar en la parte frontal de la carcasa, así fue como se llegó al diseño mostrado a continuación, el diseño fue realizado en *SolidWorks* y la manufactura fue con corte laser en acrílico, se procedió al armado de la carcasa y su posterior instalación en la cabina.

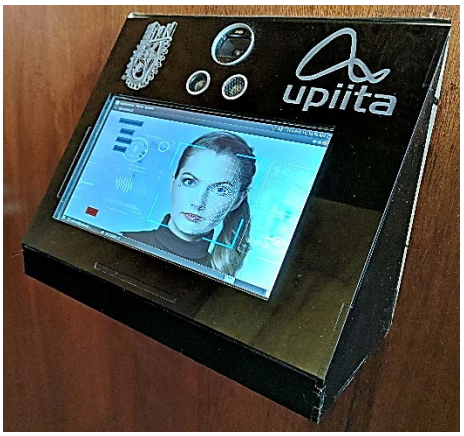


FIGURA 30 CARCASA DEL SISTEMA, VISTA DE LA INTERFAZ DE USUARIO.

9.8 Área de interés.

Como se mencionó anteriormente, se necesita una imagen cuadrada que contenga un rostro y que éste ocupe la mayor parte de la imagen, por ahora se utilizará una resolución de imagen de 640x480 pixeles, pero se puede realizar capturas de imágenes con resolución de hasta 1280x720 pixeles.

Para facilitar la extracción de características, se le indica a la persona a identificar que coloque su rostro en un recuadro de 300x280 píxeles. Posteriormente se realiza la detección de un rostro dentro de la imagen de 300x280, esto con el fin de reducir el tiempo de detección.

En la **Figura 31** se muestra el área de interés que está delimitada por un cuadrado rojo.

FIGURA 31 ÁREA DE INTERÉS.

9.8.1 Detección de rostro dentro del área de interés.

Una vez que se delimitó el área de interés, se debe hacer la captura de la imagen para su posterior procesamiento, pero de nada sirve una imagen que no contenga un rostro, es por esto que se implementó el detector en cascada con filtros de base HAAR, el cual es un método de detección de objetos efectivo propuesto por Paul Viola y Michael Jones en 2001 [19], y que esta implementado en la librería de OpenCV, de la cual utilizaremos la función *CascadeClassifier* que nos ayudara a buscar si hay algún rostro dentro del área de interés, y cuando se encuentre un rostro dentro del área de interés, se va a guardar y procesar, para hacer el reconocimiento y negar o dar el acceso.

En la **Figura 32** se muestra una imagen con nuestra área de interés y el detector de rostros funcionando, el cuadro azul significa que se ha encontrado un rostro.

FIGURA 32 DETECCIÓN DE ROSTRO.

9.8.2 Obtención de área de interés.

Una vez que se detectó un rostro dentro del área de interés, ésta se guarda para las etapas posteriores de reconocimiento facial, las imágenes que se muestran a continuación forman parte de nuestra base de datos que se utilizará durante todo el desarrollo del proyecto, para la captura de estas imágenes se ha variado la iluminación, no se ha controlado la iluminación externa, ni se ha hecho algún tipo de procesamiento digital, esto con el fin de tener la mayor cantidad de escenarios posibles en los que nuestro sistema sea capaz de funcionar como nosotros esperamos.

FIGURA 33 IMÁGENES DE BASE DE DATOS.

Cabe mencionar que estas imágenes aún tienen fondo y puede que no estén centradas, observamos esto nos trae problemas en la obtención características, ya que el rostro no siempre se encuentra en la misma posición ni ocupa el mismo espacio, es por esto que se pensó en utilizar la posición exacta del rostro y de ahí obtener una región de interés, en la que el rostro ocupe la mayor parte de la imagen dejando una cantidad mínima en la imagen que no sea rostro, en etapas posteriores se mostrará cómo fue la extracción de características de estas nuevas imágenes además de mostrar que se obtienen mejores resultados al sacar un promedio de las imágenes de la base de datos.

9.9 Normalización de imagen

Una vez que se delimitó la región de interés, se procede a normalizar las imágenes para realizar posteriormente la extracción de características.

Para lograr esto, debemos solucionar el problema de la variación de luz y obtener una imagen con iluminación homogénea, en caso de presencia de luz externa a la cabina.

Para cumplir este objetivo se puede colocar una iluminación constante enfrente del rostro de la persona, dicha luz debe estar difuminada para no causar molestia al usuario.

Se investigaron varios métodos de procesamiento de contraste en una imagen para corregir el problema de luz variante. A continuación, se muestra una imagen tomada de la base de datos (las cuales ya están en escala de grises). El rostro que se muestra tiene una de las condiciones de iluminación más extremas. Se aplicó un ajuste gama y un ajuste sigmoidal. Estos 2 métodos se utilizan para ajustar el contraste de una imagen.

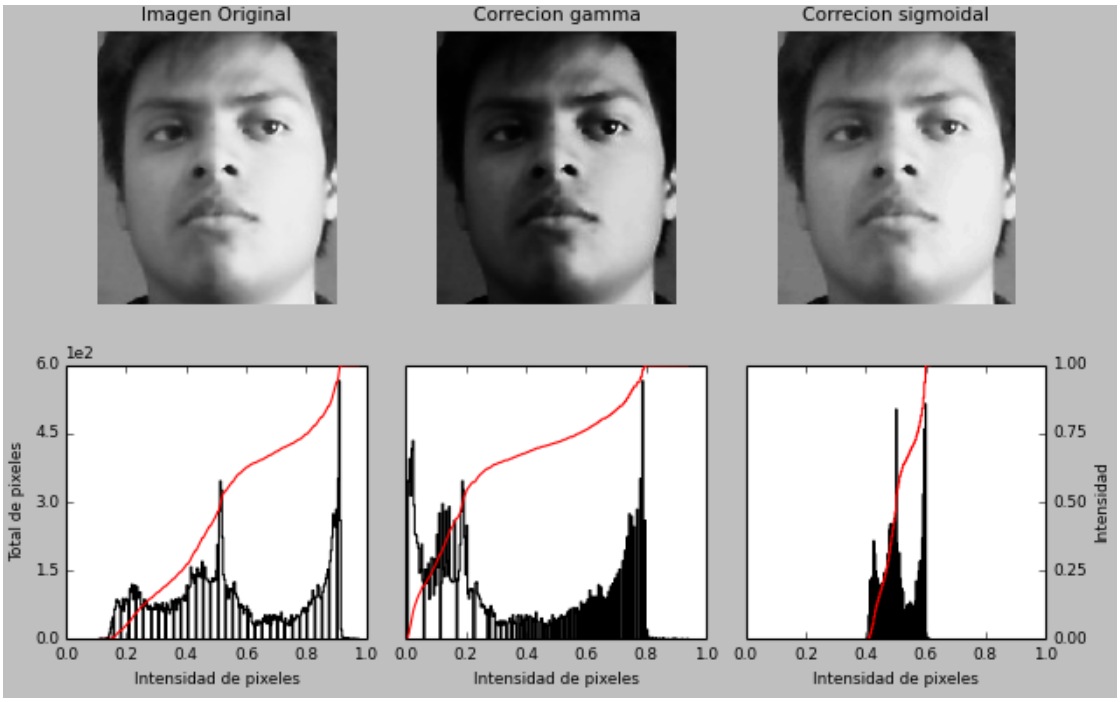


FIGURA 34 CORRECCIÓN DE CONTRASTE

Como se puede observar en el histograma, el ajuste gama distribuye la intensidad de los pixeles, al contrario del método sigmoidal, el cual concentra la intensidad de los pixeles.

A pesar de la corrección utilizada con estos 2 métodos, todavía se nota la diferencia de iluminación en los extremos del rostro. Recordemos que nuestro objetivo es tener una iluminación homogénea en todo el rostro. Para lograr esto se aplicaron otros 3 métodos, "*Contrast Streching*", "Ecualización de histograma" y "Ecualización de histograma adaptativo por límite de contraste" (CLAHE).

El mejor resultado lo da el CLAHE, pues realiza una ecualización por regiones, de tal forma que la iluminación tiende a ser homogénea.

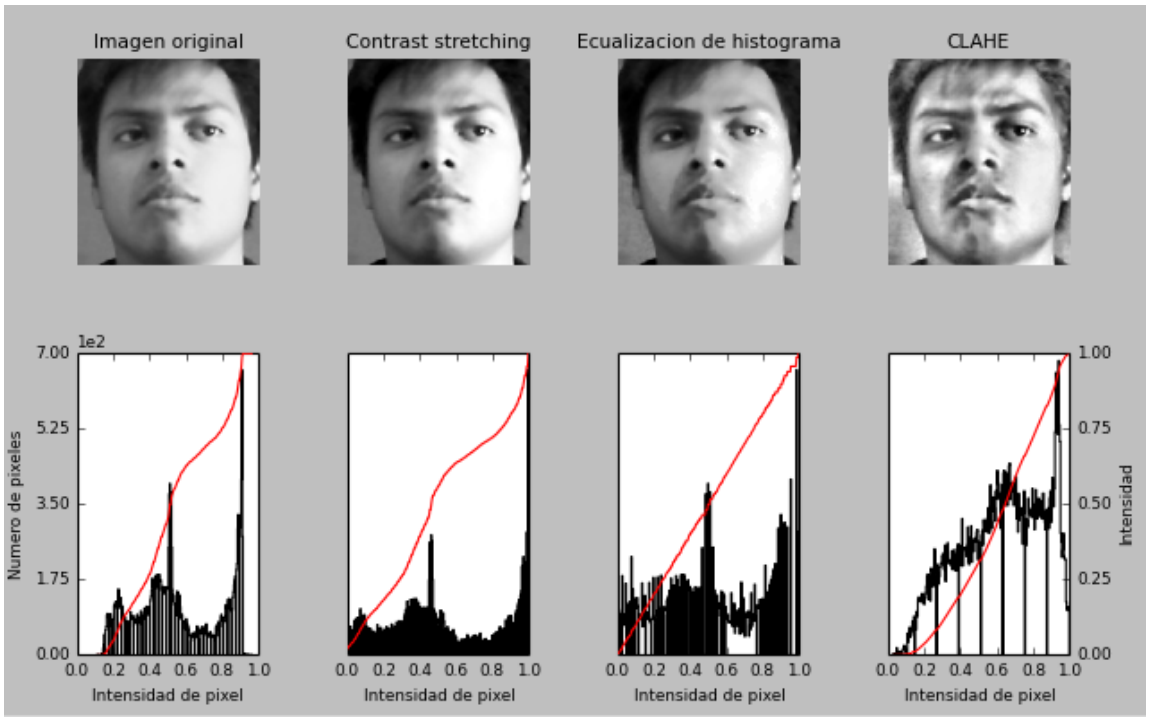


FIGURA 35 ECUALIZACIÓN DE LA IMAGEN ORIGINAL

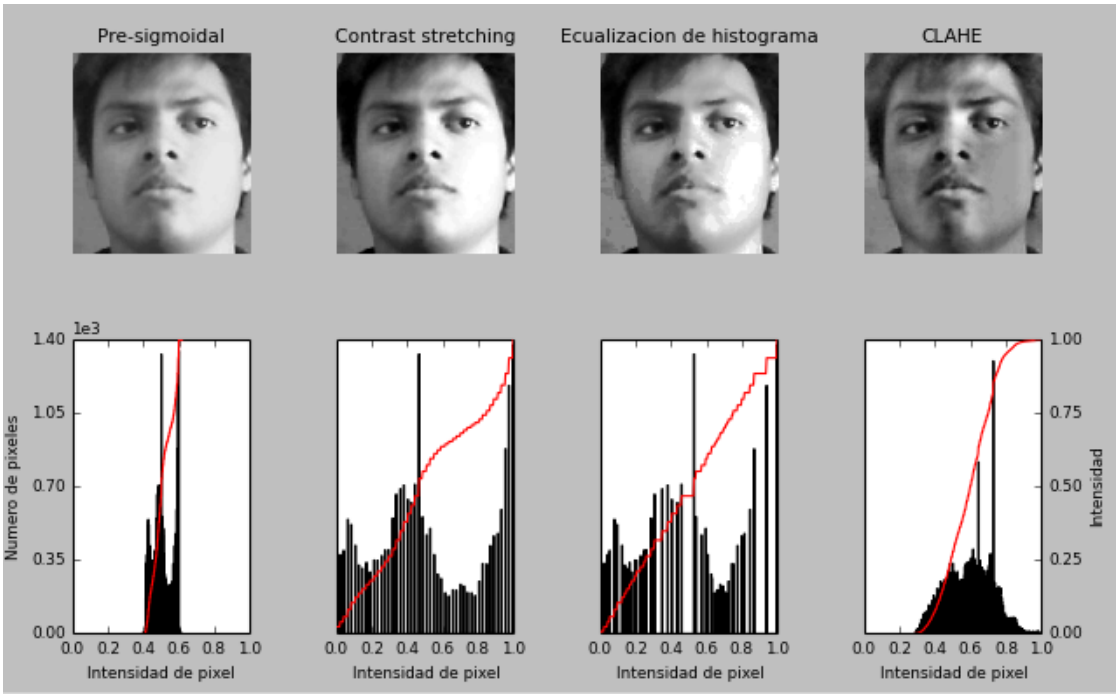


FIGURA 36 ECUALIZACIÓN DE LA IMAGEN CON CORRECCIÓN SIGMOIDAL.

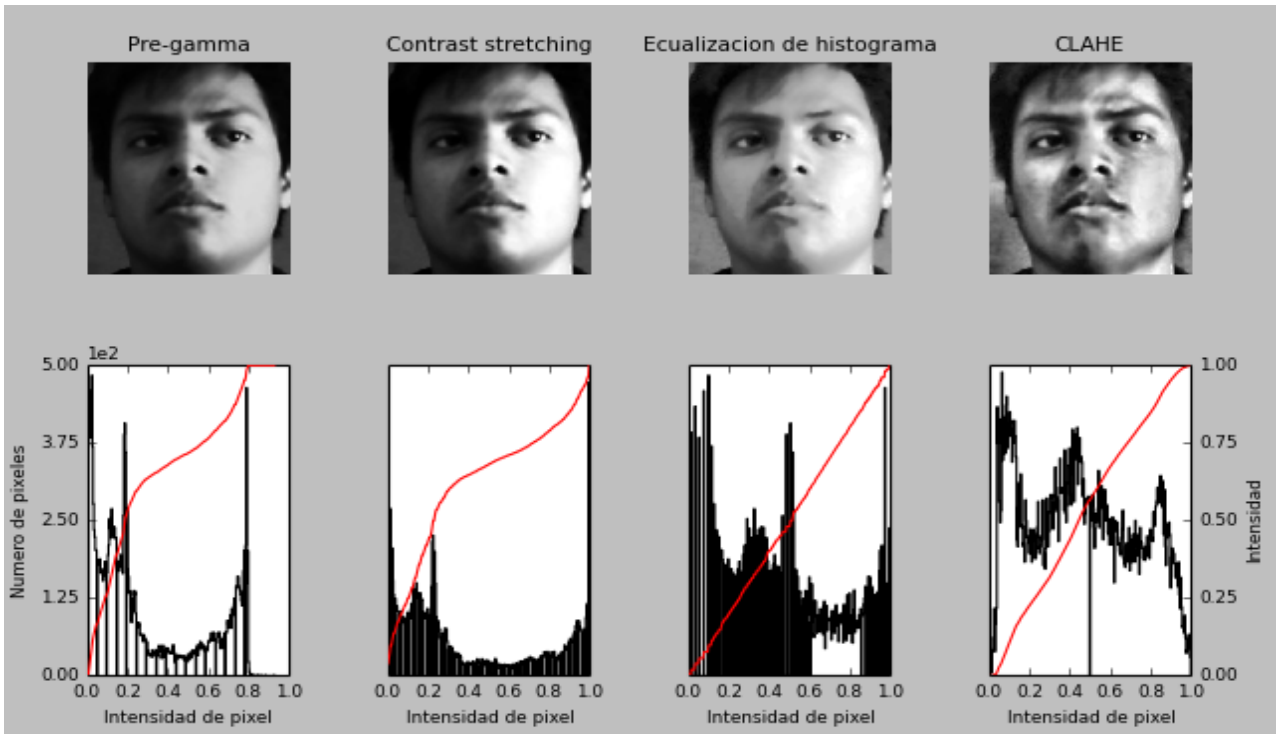


FIGURA 37 ECUALIZACIÓN DE LA IMAGEN CON CORRECCIÓN GAMMA.

9.3 Técnicas de detección de rostros genuinos

Las primeras etapas del sistema consisten en la adquisición de imágenes, la búsqueda de rostros como región de interés y el preprocesamiento. Sin embargo, dentro de estas etapas suele ocurrir lo que se conoce como “spoofing attack” o ataque de suplantación de identidad, el cual ocurre cuando una persona con el fin de falsificar el acceso de un sistema de reconocimiento facial, presenta frente a la cámara una fotografía, máscara o video de una persona con acceso autorizado.

El desarrollo de este proyecto se centra únicamente en las vulnerabilidades en 2D, es decir imágenes impresas en papel o imágenes estáticas de un dispositivo electrónico. Para la selección de un método para solucionar esta problemática se analizaron los métodos basados en movimiento, análisis de textura y basados en hardware.

TABLA 6 RESUMEN DE ANÁLISIS DE MÉTODOS PARA LA DETECCIÓN DE "SPOOFING ATTACKS" [47]

Método	Ejemplos presentados en artículos	Características
Métodos basados en análisis de movimiento	Detección de parpadeo Detección de movimiento analizando la relación entre movimiento del rostro y el movimiento del entorno.	Alto coste computacional Buena capacidad de respuesta. Fácilmente eludido por falsos movimientos
Métodos basados en análisis de textura	Análisis de textura empleando un modelo Lambertiano Análisis de textura usando LBP	Respuesta rápida Bajo coste computacional Baja eficiencia ante imágenes de buena calidad
Métodos basados en Hardware	Combinación de detección de rostro y voz. Empleo de cámara térmica Empleo de cámaras 3D	Buena capacidad de detección. Requiere elementos de hardware extra. Eleva el costo del sistema.

9.4 Selección del método a emplear para la detección de rostro genuino

Hasta la fecha no existe un método que se considere como el mejor para resolver la vulnerabilidad de "spoofing face". Este problema ha sido identificado en artículos desde algunos años y se ha tratado reducir la vulnerabilidad empleando distintos métodos o combinación de ellos, nuestro proyecto solo se centra en solventar el problema de vulnerabilidades en 2D.

En base a la **Tabla 6** se decidió emplear una técnica basada en el análisis de movimiento que consiste en detectar parpadeo, mediante un clasificador en cascada que fue entrenada con imágenes de parpadeos disponible en el repositorio de OpenCV.

También se ha de emplear este método en conjunto con un método basado en Hardware que consiste en medir la distancia que existe entre el rostro y la cámara para así descartar imágenes de teléfonos móviles que se encuentren frente a la cámara.

9.4.1 Detección de rostro genuino mediante detección de parpadeo

Una vez implementado el clasificador para detección de parpadeo como se muestra en la **Figura 38,** se ha de mostrar un cuadro verde que encierra el parpadeo mediante la interfaz gráfica al usuario cuando esto sucede, esto con la finalidad de que la persona que se somete a reconocimiento pueda interactuar con el sistema, sabiendo que para que el acceso sea otorgado debe parpadear al menos una vez durante el proceso. Esta es una desventaja en cuanto al tiempo del proceso de control de acceso, ya que el proceso de reconocimiento es mucho rápido en comparación con el promedio de tiempo que tarda una persona en parpadear. Pero a su vez, el parpadeo proporciona certeza de que se trata de un rostro genuino.

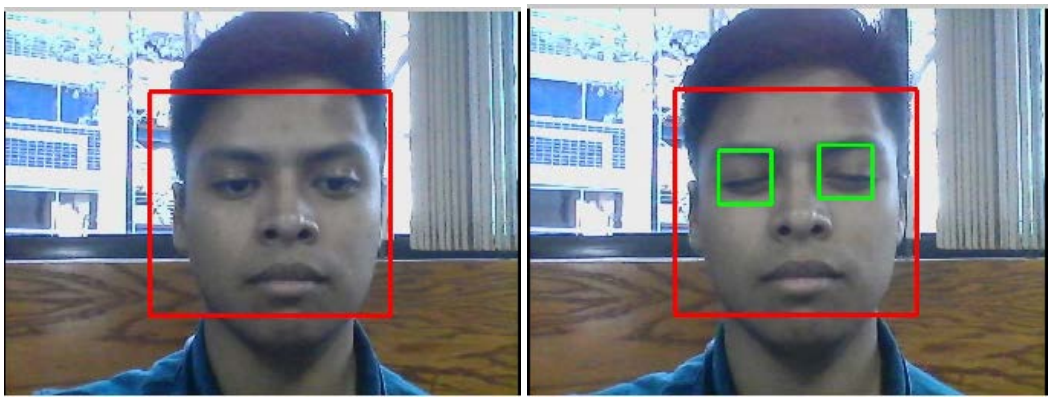


FIGURA 38 DETECCIÓN DE PARPADEO

Se realizaron pruebas donde al presentar la fotografia de un usuario del sistema en un telefono móvil, el detector de parpadeo funciona correctamente ya que no se detecta parpaedo alguno.

FIGURA 39 DETECCIÓN DE ROSTRO EN UN TELÉFONO MÓVIL, DONDE NO SE DETECTA PARPADEO

Sucede algo similar cuando se presentan fotografías impresas en papel las cuales se colocaron frente a la cámara, se agitaron, rotaron y no se detectó parpadeo.

FIGURA 40 DETECCIÓN DE ROSTRO DE UNA IMAGEN IMPRESA, DONDE NO SE DETECTA PARPADEO

9.4.2 Detección de rostro genuino mediante sensor de distancia

Con el fin de combinar un método basado en movimiento (detección de parpadeo) con un método basado en Hardware, se optó por utilizar un sensor ultrasónico, ya que con este podemos calcular la distancia a la que se encuentra un objeto plano, esto con la finalidad de detectar teléfonos móviles o fotografías que estén cerca de la cámara. Al principio se consideró emplear un sensor térmico para detectar la temperatura de los objetos frente a la cámara, y descartar aquellos que no correspondieran con la temperatura promedio de un cuerpo humano, sin embargo, al emplear sensores de este tipo suelen haber problemas por ruido generado por los cambios de temperatura en el ambiente, y su elevado costo. Por lo tanto, se optó por emplear un sensor ultrasónico HC-SR04 que es un sensor económico y que no requiere de filtros o acondicionamiento de señal adicionales para su empleo.

El sensor ultrasónico es usado de tal manera que cualquier objeto plano que se encuentre a menos de 30 cm, como puede ser una hoja de papel, o la pantalla de un teléfono móvil, lleve al sistema a descartar cualquier posibilidad de otorgar el acceso. Esta información es enviada desde la tarjeta de programación a la microcomputadora a través de la publicación de un tópico, implementando ROS en la tarjeta de programación.

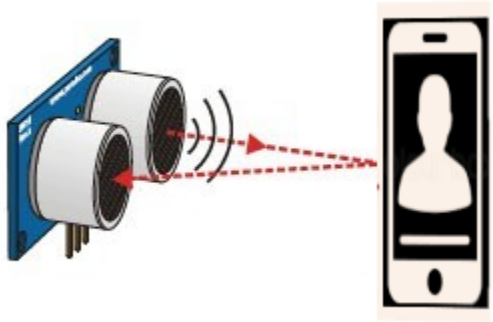


FIGURA 41 SENSOR ULTRASÓNICO HC-SR04 PARA DETECTAR OBJETOS PLANOS CERCANOS

El sensor ultrasónico presenta una función doble, al identificar presencia delante de la cámara, se enciende el panel, y una vez que detecta la ausencia de un objeto o persona frente a la cámara, apagará el panel de iluminación después de 15 segundos con la finalidad de ahorrar energía eléctrica.

9.5 Esquema del sistema en ROS

El sistema desarrollado ha sido programado mediante uso de lenguaje Python en donde se han agregado librerías que permiten su implementación con ROS, el cual es un sistema basado en una arquitectura de grafos donde el procesamiento toma lugar en nodos. Un "nodo" es un ejecutable en código que permite enviar, recibir, y multiplexar mensajes.

Cada nodo puede publicar mensajes a través de un "tópico" el cual es un canal de información o mensajes, éste puede transportar distintos tipos de datos tales como imágenes, valores numéricos, caracteres, estados lógicos, entre otros. Para que un nodo pueda enviar o recibir un mensaje, basta con "subscribirse" a un tópico o "publicar" en él.

En la **figura 43** se muestra un gráfico generado por la herramienta "rqt_graph" de ROS en el cual se visualiza el funcionamiento del sistema a través de los nodos y tópicos del sistema. En este gráfico los nodos estan representados por un elipse y los tópicos por un recuadro, dentro de cada elemento se encuentra su nombre correspondiente.

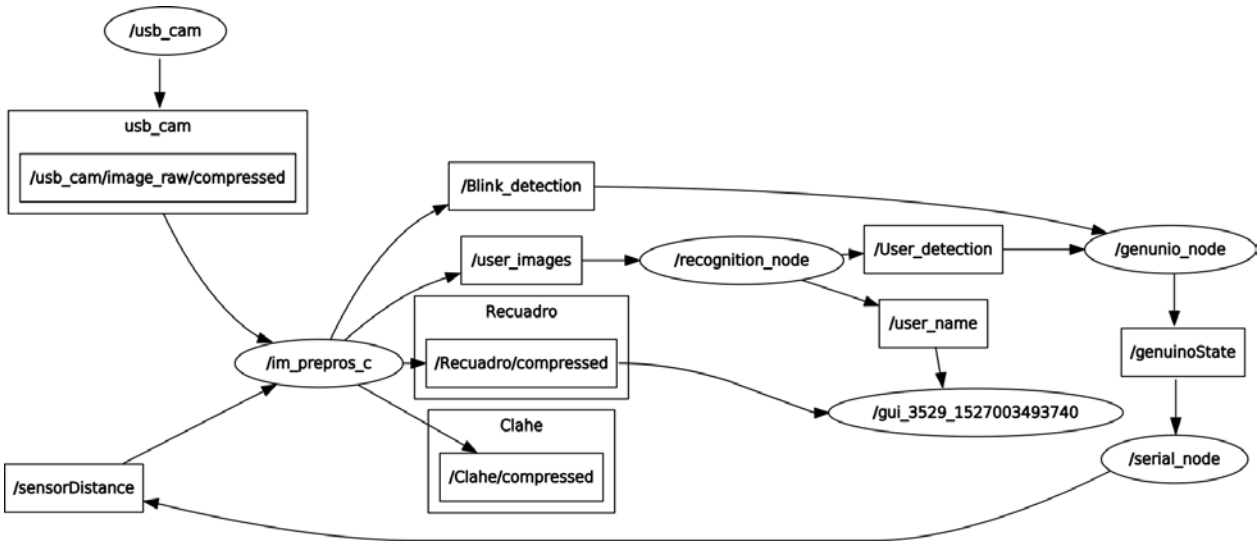


FIGURA 42 GRÁFICO DEL SISTEMA EN ROS

El proceso comienza en el nodo "/usb_cam", en el cual se realiza la adquisicion de imágenes, y se ajustan parametros tales como el tamaño de la imagen y el número de cuadros por segundo. Éste publica las imágenes obtenidas en la cámara de manera continua en un tópico al que se subscribirá el nodo de preprocesamiento.

La etapa de preprocesamiento está implementada en el nodo "/im_prepros_c". En este nodo se lleva a cabo la transformación de datos de ROS en imágenes, posteriormente se recorta la imagen adquirida. Usando hrramientas de OpenCV se dibuja un recuadro que permite la reducción del área de interés. Enseguida se aplican clasificadores en cascada para la deteccion de rostro en la imagen y detección de parpadeo.

Se convierten las imágenes a escala de grises y se les aplica el método de ecualización adaptativo "CLAHE" el cual permite contrarrestar defectos originados por la iluminación.

Despúes, se guarda un conjunto de 10 imágenes ya normalizadas que servirá para la etapa de clasificación. Esto con el fin de asegurar que la persona que se va a reconcer sea clasificada correctamente, es decir, se clasifican 10 mágenes por cada intento. Se publica un estado lógico para saber que esta tarea ya se ha realizado y se reinicie todo el proceso para cada vez que un usuario solicite su reconocimiento.

En cuanto a la identificación de rostro genuino, se realiza un monitoreo de la distancia que hay frente a la cámara. Si existe un objeto muy cerca de la cámara, no se realiza el preprocesamiento ni las etapas consecuentes. La información obtenida en este nodo se envia a traves de tópicos a otros nodos.

El nodo "/recognition_node" está asociado a la etapa reconocimiento facial, se toman las 10 imágenes generadas en la etapa de preprocesamiento y se realiza la extracción de caracteristicas, consecuentemente se aplica el mecanismo de clasificación y se genera una respuesta según el resultado del clasificador. Se publica en un nodo el nombre del usuario si es que este coincide con alguno de los rostros almacenados en la base de datos, en caso contrario se publica una leyenda de "Usuario no registrado".

El nodo "genuino_node" corresponde a un programa que se subscribe a tópicos para obtener información acerca del estado del parapedo, y del reconocimiento facial, es decir el sistema accionará el mecanismo de la puerta si se cumplen las siguiente dos condiciones:

- Se ha detectado al menos un parpadeo durante el proceso de reconocimiento, por lo tanto se trata de un rostro genuino y no de una fotografía.
- El usuario que solicita el acceso, corresponde a una persona registrada en la base de datos y por lo tanto se le ha otorgado el acceso.

9.10 Extracción de características

Una vez finalizado el preprocesamiento de las imágenes, éstas son sometidas a un proceso donde se extraen datos únicos y representativos de los rostros. Esta información se ve reflejada en un conjunto de datos que serán indispensables para la comparación con un rostro de prueba. La etapa de extracción de características es vital en el reconocimiento facial ya que, si esta etapa no se realiza adecuadamente, aun empleando el mejor clasificador no se obtendrán resultados apropiados.

En la extracción de características se obtiene como resultado la representación de una imagen en un conjunto de datos. Las buenas representaciones son aquellas que por un lado minimizan las diferencias entre imágenes de una misma persona y por el otro lado maximizan la diferencia entre imágenes de personas distintas, se espera también que la extracción de características sea rápida y compacta.

Existen muchos enfoques sobre la clasificación de los métodos de extracción de características. Uno de ellos divide métodos en dos tipos; los métodos holísticos (métodos basados en apariencia) y los métodos basados en características locales.

En el primer tipo de métodos, la imagen de la cara completa es usada como una entrada para la operación de reconocimiento, tal como ocurre con PCA [47], cada imagen del rostro se representa como un punto en un subespacio vectorial, de forma que la comparación entre la imagen a prueba y las de referencia se realiza en el subespacio vectorial de caras

[1].En el segundo tipo de métodos se extraen las características según la distribución de rasgos faciales tales como la ubicación y distancia de los ojos, nariz y la boca.

En este trabajo se emplearán dos técnicas de métodos basados en apariencias, el primero es PCA, y el segundo es LDA, ambos mencionados anteriormente.

9.10.1 Extracción de características por análisis de componentes principales PCA o "Eigenfaces"

El primer método empleado es PCA. Se utilizó la librería de *scikit-learn* [40] para realizar la extracción de características. La documentación de esta librería indica que el número máximo de componentes principales que se pueden extraer es el siguiente:

$$n_componentes = min(n_características, n_muestras)$$

En el caso de este proyecto, se cuenta con una base de datos de 1190 imágenes de 140x160 píxeles; cada imagen cuenta con 22400 píxeles o características. Antes de realizar el proceso de PCA, se dividió y se mezcló la base de datos, esto con el fin de tener un mejor rendimiento en el proceso de clasificación. Se usaron 20% de las imágenes para pruebas y 80% de ellas para el entrenamiento, escogidas de manera aleatoria. De tal forma que el número máximo de componentes principales es 952.

Una de las principales ventajas de utilizar PCA es la reducción de dimensión de los datos. PCA crea un nuevo sistema coordenado para cada componente principal, donde se minimiza la distancia de los puntos al nuevo eje x' y se minimiza la pérdida de información al proyectar cada punto en el eje x'.

Cabe destacar que la componente principal cuenta con el mayor porcentaje de varianza, es decir, el individuo más distinto con respecto al promedio, será el más fácil de clasificar.

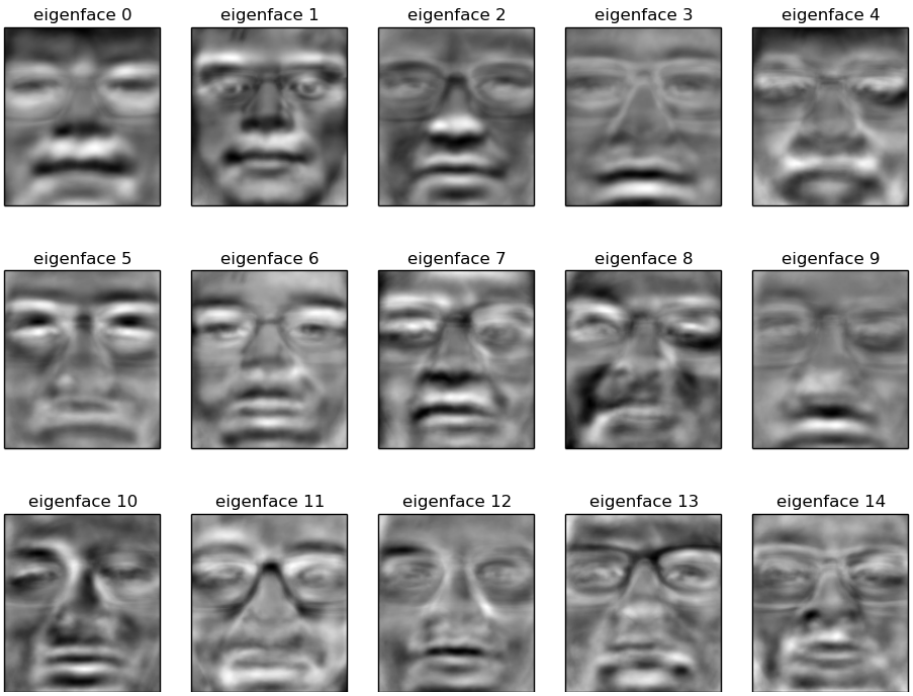


FIGURA 43 RECONSTRUCCIÓN DE ROSTROS A PARTIR DE COMPONENTES PROPORCIONALES

En la **Figura 43**, se muestra la reconstrucción de los primeros 15 componentes principales. A estos componentes también se les conoce como *eigenfaces*.

Se puede notar que en los primeros componentes o *eigenfaces,* es incluso posible diferenciar de qué persona se trata. Cada *eigenface* corresponde a la reconstrucción de cada componente principal. Es notorio que la primera componente principal es el que tiene más varianza. Esto es porque un individuo de la base de datos tiene cabello sobre el rostro, además de que es la única mujer en la base de datos, por lo que es el más distinto respecto a la imagen promedio y por esta razón es el que tiene más varianza.

En la **Figura 44**, se muestra la reconstrucción de la imagen promedio de las 80 imágenes de entrenamiento.

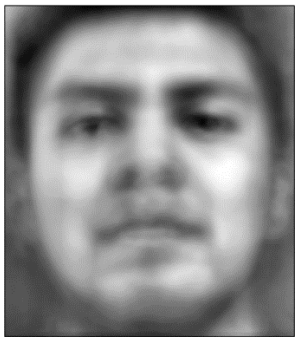


FIGURA 44 IMAGEN PROMEDIO DE LA BASE DE DATOS

Si se agrega una nueva persona, a la base de datos, que tenga rasgos similares a la imagen promedio, esta persona será muy difícil de clasificar, ya que sus componentes principales serán parecidos.

También se realizó la imagen promedio de nuestra base de datos con los tres métodos realizados en el preprocesamiento de imagen. Estos tres métodos son:

A. CLAHE sin corrección de contraste

B. CLAHE con corrección sigmoidal.

C. CLAHE con corrección gamma.

FIGURA 45 A) B) C)

Hay que destacar que para saber cuál de los 3 métodos de procesamiento de imagen es mejor, se tiene que evaluar con la precisión del clasificador, la cuál es la siguiente etapa.

9.10.2 Extracción de características por Análisis lineal discriminante (LDA) o "Fisherfaces".

Se eligió el método *Fisherfaces* o LDA *(Linear Discriminant Analisis)* como segundo método de extracción de características, ya que ofrece algunas ventajas interesantes para nuestro proyecto. En contraste con *Eigenfaces*, el método *Fisherfaces* busca maximizar la varianza de las muestras entre clases (entre personas) y minimizarla entre muestras de la misma clase (de la misma persona). Con esto se logra obtener mejores resultados en caso de que haya variaciones de la iluminación y expresión respecto de las imágenes de entrenamiento. Por otro lado, debemos hacer una base de datos más robusta que tenga imágenes de rostros tomadas con iluminación variable, así como diferentes poses y gestos.

De esta manera, la descomposición en *fisherfaces* interpola o extrapola las demás condiciones de pose o iluminación que se pueden presentar aparte de las ya tomadas. El proceso de reconocimiento es básicamente el mismo que en el PCA: primero, la imagen es proyectada al espacio formado por las imágenes de entrenamiento, para a continuación buscar la imagen más semejante por medio de la SVM, esto se realiza 10 veces cada que se prueba el sistema, con el fin de minimizar errores en algunas de las capturas en donde el rostro pudo haberse movido o la iluminación no haya sido correcta, o la exposición automática de la cámara aún no se haya estabilizado, teniendo una tasa reconocimiento más estable. Por otro lado, lo que diferencia ambas técnicas es la forma de calcular el espacio donde están proyectadas las imágenes de entrenamiento. [48]

Para calcular el espacio, se buscan dos matrices de covarianza, una inter-clase que corresponde a las diferentes imágenes de una misma persona, y una extra-clase que corresponde a las imágenes de diferentes personas. En la **Figura 46** podemos ver un ejemplo de dos formas distintas de proyectar los mismos puntos sobre una línea, una buscando minimizar la distancia entre puntos de una misma clase, y otra buscando maximizar la distancia entre puntos de distinta clase.

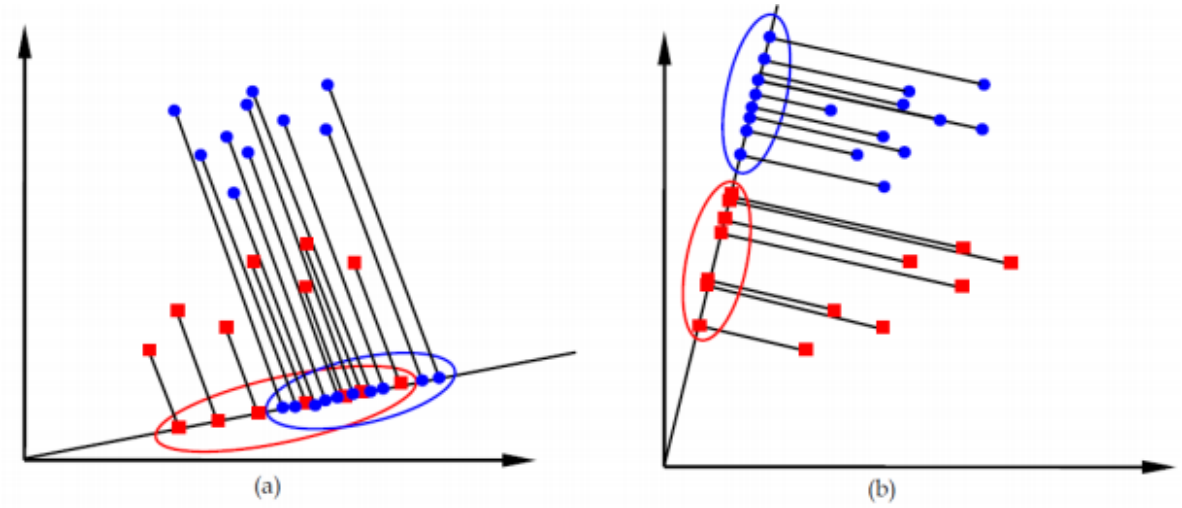


FIGURA 46 PUNTOS PROYECTADOS A UNA LÍNEA. (A) MINIMIZANDO LA DISTANCIA (B) MAXIMIZANDO LA DISTANCIA [48]

Para la implementación de este método de extracción de características se va a utilizar la librería de scikit-learn [40], como en el método anterior, esto para facilitar su implementación

y compatibilidad con el código antes generado para extracción de características y clasificación.

9.11 Clasificación

En la etapa de clasificación se decidió utilizar el método SVM ya que se han logrado resultados satisfactorios para el reconocimiento de rostros. El método SVM ha sido incluso más preciso que otros métodos de clasificación como MLP (*Multilayer perceptron*) [48].

Uno de los objetivos particulares es el siguiente:

- Evaluar el funcionamiento global del sistema con el uso de 2 técnicas para la extracción de características que permitan la identificación genuina de rostro, con el fin de garantizar una eficiencia de al menos 70%. Se implementará solo la técnica que presente el mayor índice de identificación correcta.

Cuando se refiere a eficiencia, se refiere a lo siguiente:

$$eficiencia = \frac{no.\ de\ intentos\ clasificados\ correctamente}{no.\ total\ de\ intentos}$$

Sin embargo, tal vez esta no sea la mejor forma de evaluar el funcionamiento del sistema. Por ejemplo, supongamos dos escenarios:

Primer escenario. Se realizan 10 intentos, 7 de ellos corresponden a personas registradas en la base de datos y fueron clasificadas correctamente. Los 3 intentos restantes corresponden a personas que no están en la base de datos, pero se les otorgó el acceso.

Segundo escenario. Se realizan 10 intentos, 7 de ellos corresponden a personas no registradas en la base de datos y fueron clasificadas correctamente. Los 3 intentos restantes corresponden a personas que sí están registradas, pero no se les otorgó el acceso.

Estos dos escenarios son totalmente distintos; sin embargo, no tienen el mismo impacto en el funcionamiento del sistema. ¿Qué escenario es peor, otorgar el acceso a personas no registradas o negar el acceso a personas registradas?

Al ser un sistema de seguridad, es preferible que una persona que sí está registrada realice varios intentos, a que una persona no registrada obtenga el acceso. Entonces, lo que se requiere es un mejor funcionamiento para clasificar correctamente a las personas no registradas.

Para evaluar el funcionamiento global del sistema es mejor referirse a dos términos distintos, “precisión y *recall*”. Estos se refieren al funcionamiento de cada clase.

$$recall = \frac{VP}{VP + FN} \quad donde: \quad VP = Verdadero\ Positivo, FN = Falso\ Negativo$$

El *recall* se refiere al porcentaje en el que algún elemento de una clase específica fue clasificado como otra.

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

La precisión se refiere al porcentaje en el que elementos de otras clases fueron confundidos con una clase específica.

En términos de este proyecto:

Supongamos la clase Sophia, un usuario registrado en la base de datos.

El *recall* corresponde a el porcentaje en el que Sophia fue confundida con otra clase.

Precisión corresponde a el porcentaje en el que otras clases fueron confundidas con Sophia.

Lo ideal es que tanto la precisión como el *recall* sean cercanos a 1; sin embargo, en este proyecto la clase "Usuario no registrado", tiene el mayor impacto en el funcionamiento. Es más importante que el *recall* sea cercano a 1, es decir, que no se dé acceso a personas no registradas.

9.11.1 Clasificación con PCA como método de extracción de características.

Después de aplicar PCA se redujo de 22400 a 80 componentes cada imagen, por lo que se entrenó la SVM con 952 imágenes de 80 componentes. La suma de la varianza de estos 80 componentes corresponde a 0.85. Para que la suma de la varianza de los componentes sea 1, se debe obtener el número máximo de componentes, que corresponde en este caso al número total de imágenes en la base de datos (1190 imágenes).

Se utilizó la librería de *scikit-learn* con un *kernel* RBF (*Radial Basis Function*). También se utilizó una función llamada $GridSearch()$, la cual encuentra el mejor parámetro de *kernel* para obtener el mayor porcentaje de precisión.

En la **figura 41** se muestra el porcentaje normalizado de *precision*, *recall* y *score* correspondiente a cada clase. *Score* corresponde al promedio entre *precision* y *recall*. El *support* se refiere al número de imágenes tomadas de la base de datos para realizar pruebas.

	precision	recall	f1-score	support
Gerardo Fosado	1.00	0.93	0.97	15
Naham Tornes	1.00	1.00	1.00	14
Vanessa Lopez	1.00	1.00	1.00	13
Olaf Perez	0.94	1.00	0.97	15
Octavio Montes	1.00	1.00	1.00	16
Adrian Blas	1.00	1.00	1.00	9
Andrea Ponce	1.00	1.00	1.00	14
Eduardo Angel	1.00	1.00	1.00	11
Eduardo Cruz	1.00	1.00	1.00	12
Abraham Rodrioguez	1.00	1.00	1.00	14
Pablo Torres	1.00	1.00	1.00	23
Angel Villalobos	1.00	1.00	1.00	16
Alejandro Pena	1.00	1.00	1.00	15
Ricardo Elizalde	1.00	1.00	1.00	11
Luis Forntanel	1.00	1.00	1.00	9
Hector Martinez	1.00	1.00	1.00	17
Daniel Miranda	1.00	1.00	1.00	14
avg / total	1.00	1.00	1.00	238

FIGURA 47 REPORTE DE CLASIFICACIÓN TEST PCA.

Cabe destacar que, a pesar del buen rendimiento, esta prueba no contempla a usuarios no registrados. Si se realizan pruebas con imágenes de usuarios que no se encuentran en la base de datos, el modelo de entrenamiento clasifica a estos usuarios según el mayor parecido con los usuarios existentes.

El clasificador SVM, tiene un método de clase llamado "*predict_proba*", el cual regresa la probabilidad que una imagen corresponda a cada una de las clases. La clase elegida es la que tiene la probabilidad más alta.

La matriz de confusión queda de la siguiente forma:

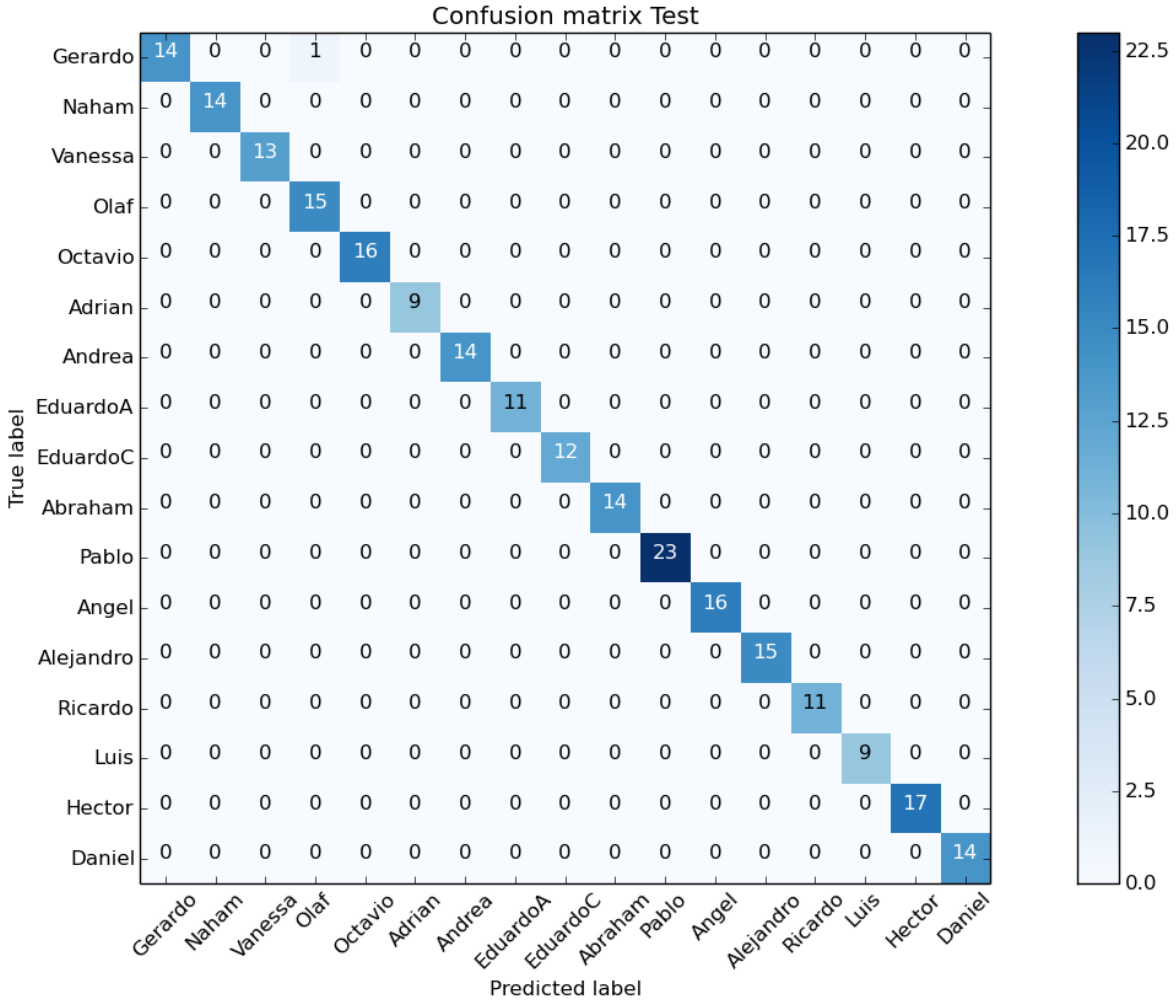


FIGURA 48 MATRIZ DE CONFUSIÓN

Como se puede observar en el reporte y en la matriz, sólo 1 imagen (de un total de 238 imágenes) fue clasificada de manera incorrecta. Es por ello que se decidió por entrenar con un total de 80 componentes.

Por otra parte, se plantearon tres algoritmos de preprocesamiento de imagen:

1. CLAHE sin corrección de contraste
2. CLAHE con corrección sigmoidal.
3. CLAHE con corrección gamma.

Los resultados previamente presentados, se basan en un preprocesamiento usando CLAHE sin corrección de contraste. Debido al alto porcentaje de imágenes correctamente clasificadas, se decidió optar por este método de procesamiento.

En la **Figura 49** se muetran las 20 imágenes de prueba clasificadas de manera correcta.

FIGURA 49 TEST CON UN CONJUNTO DE 20 IMÁGENES ALEATORIAS

9.11.2 Clasificación con LDA como método de extracción de características.

Después de aplicar LDA se obtuvieron 16 componentes, ya que el número máximo de componentes en LDA es $n_{clases} - 1$. La suma de la varianza total de estos componentes corresponde a 1.0, a diferencia de PCA, donde la suma de la varianza de los 80 componentes seleccionados corresponde a 0.85.

Básicamente es el mismo procedimiento que se realizó con el método PCA.

	precision	recall	f1-score	support
Gerardo Fosado	1.00	0.93	0.97	15
Naham Tornes	1.00	0.93	0.96	14
Vanessa Lopez	1.00	1.00	1.00	13
Olaf Perez	1.00	1.00	1.00	15
Octavio Montes	1.00	1.00	1.00	16
Adrian Blas	1.00	1.00	1.00	9
Andrea Ponce	1.00	1.00	1.00	14
Eduardo Angel	1.00	1.00	1.00	11
Eduardo Cruz	1.00	1.00	1.00	12
Abraham Rodrioguez	1.00	1.00	1.00	14
Pablo Torres	1.00	1.00	1.00	23
Angel Villalobos	1.00	1.00	1.00	16
Alejandro Pena	0.94	1.00	0.97	15
Ricardo Elizalde	1.00	1.00	1.00	11
Luis Forntanel	1.00	1.00	1.00	9
Hector Martinez	0.94	1.00	0.97	17
Daniel Miranda	1.00	1.00	1.00	14
avg / total	0.99	0.99	0.99	238

FIGURA 50 REPORTE DE CLASIFICACIÓN TEST LDA

La principal diferencia en el algoritmo es que LDA al ser un método de extracción supervisado, necesita tanto los pixeles de la imagen (componentes X) cómo la clase a la que pertenece (nombre del usuario Y). PCA sólo necesita los pixeles.

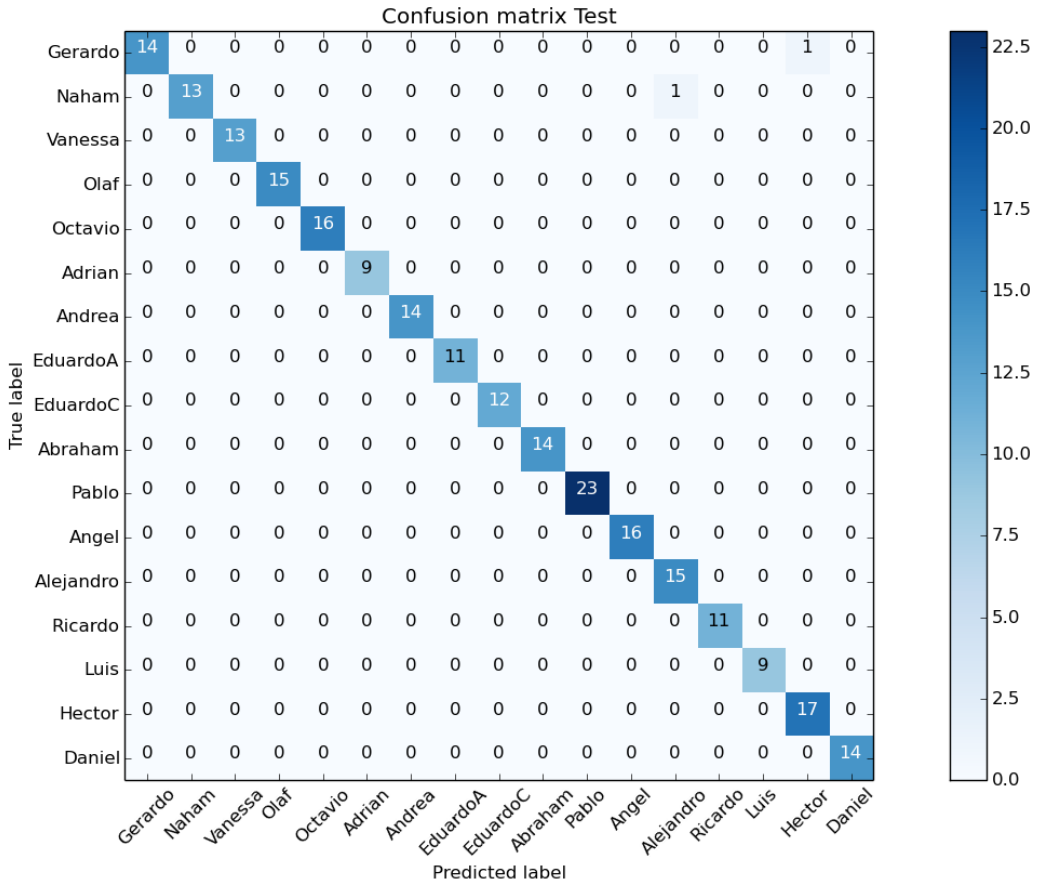


FIGURA 51 TEST CON UN CONJUNTO DE 20 IMÁGENES ALEATORIAS USANDO LDA

FIGURA 52 IMÁGENES ALEATORIAS CLASIFICADAS USANDO LDA

En esta etapa, ambos métodos arrojaron resultados similares y con buen desempeño. Para comparar de mejor forma el rendimiento de éstos, es necesario realizar diversas pruebas con usuarios registrados y no registrados. Dichas pruebas se muestran en la sección análisis de resultados.

9.12 Desarrollo de interfaz gráfica de usuario (GUI).

Para el desarrollo de la interfaz gráfica se tienen ciertos requerimientos en el proyecto. El primero es que se pueda mostrar la información de uno o más tópicos de ROS, es decir, que exista una comunicación entre la interfaz y el ambiente de ROS, entre ellos un tópico de tipo imagen que va a mostrar lo que está viendo la cámara y alguna referencia para que el usuario se posicione. Otro punto importante es que la curva de aprendizaje sea rápida, ya que no se tiene mucho tiempo para el desarrollo de esta interfaz. Esos son los dos puntos principales que se consideraron para elegir la herramienta de desarrollo de la interfaz.

Se seleccionó la librería Tkinter, incluida por defecto en Python, debido a que el desarrollo de GUI´s con Tkinter es claro y tiene una curva de aprendizaje rápida. Su entorno de trabajo es orientado a objetos, como casi todas las herramientas de desarrollo de interfaces gráficas, además de que con un poco de entendimiento del funcionamiento de ROS y de la interfaz de usuario desarrollada con Tkinter, se puede crear una comunicación entre ROS y la GUI, que para nosotros es un punto importante para el proyecto.

En la **Figura 53** Vista de la GUI se puede observar la primera versión de la GUI, parece muy simple pero ya cuenta con una comunicación con ROS, al suscribirse a tópicos, además de iniciar algunos lanzadores desde la interfaz, esto facilita la forma en que se agrega un nuevo usuario a la base de datos.

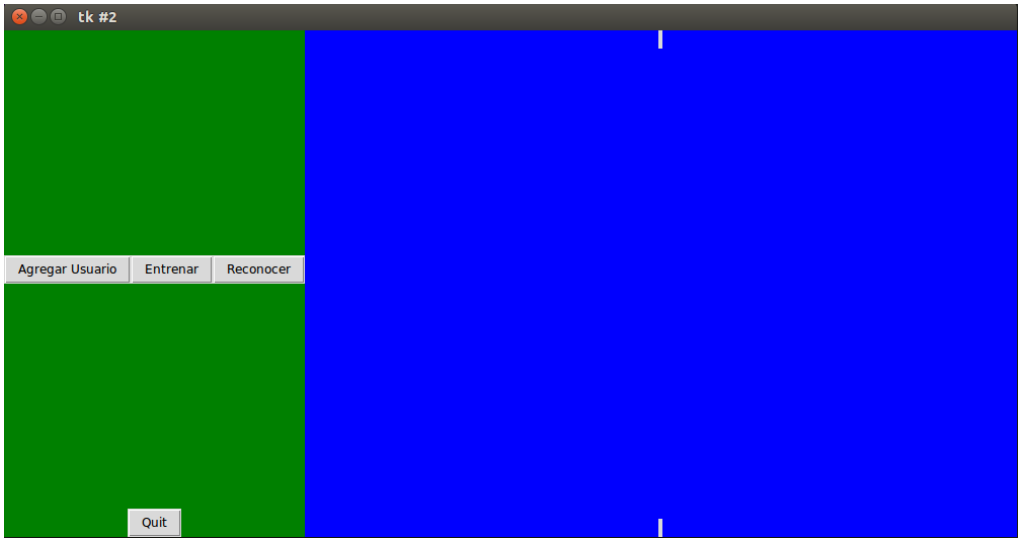


FIGURA 53 VISTA DE LA GUI

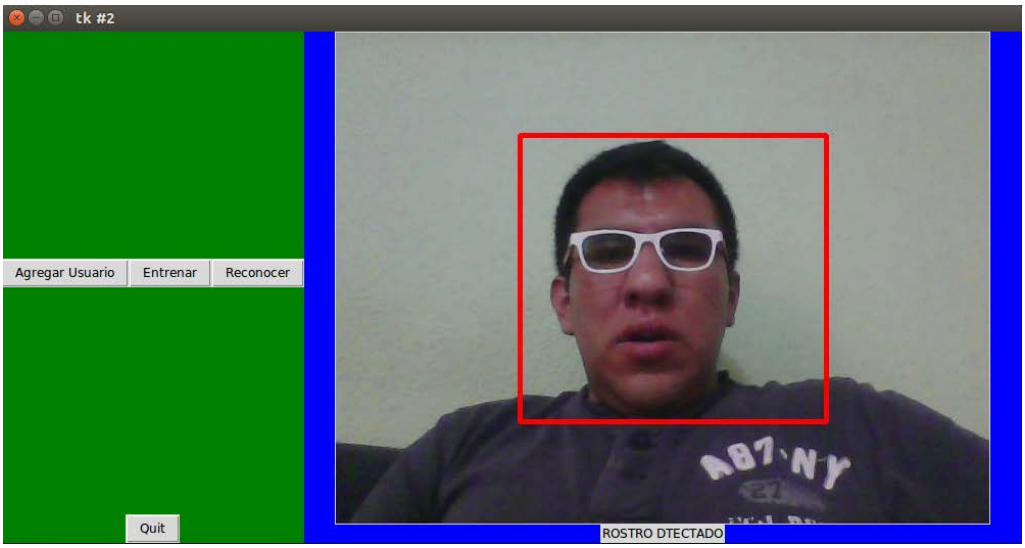


FIGURA 54 GUI FUNCIONANDO UNA VEZ QUE SE PRESIONÓ EL BOTÓN "RECONOCER"

En las siguientes figuras se muestra la versión final de la interfaz gráfica de usuario. Se crearon varias versiones. Una versión general en la que se tienen 5 botones: agregar usuario, entrenar, reconocer, detener y salir. Cada uno tiene una función específica, los cuales activan o desactivan elementos de la interfaz, además de realizar acciones relacionadas con el funcionamiento del sistema.

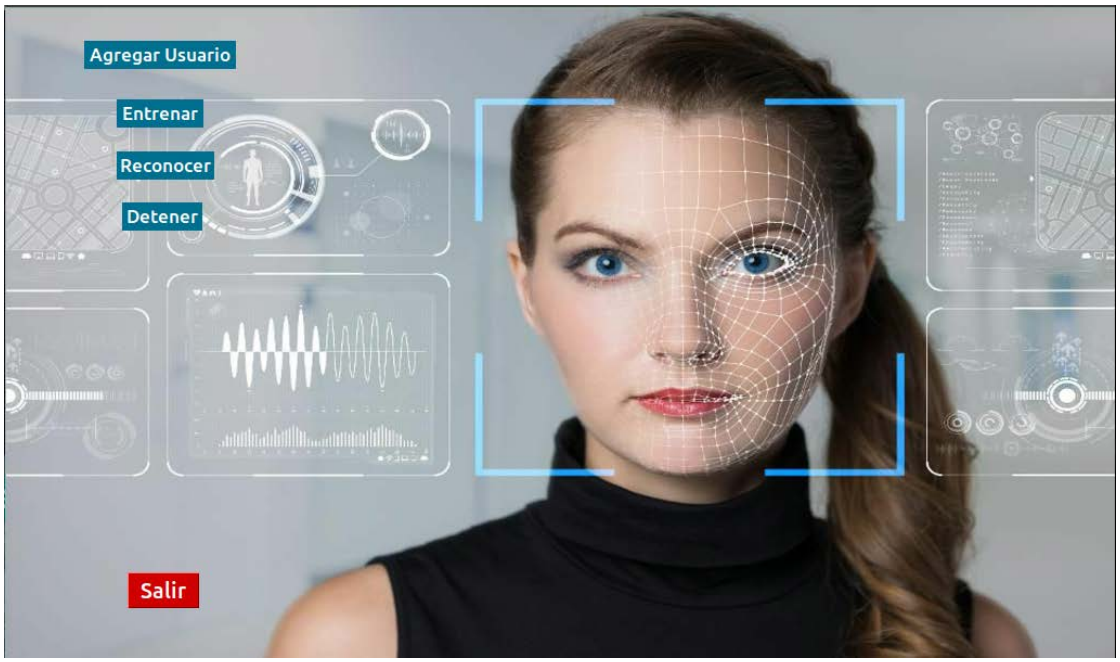


FIGURA 55 VISTA FINAL DE LA GUI

Para el desarrollo de la versión final se utilizó una herramienta gráfica para el desarrollo de GUI´s llamada Page, que usa la librería Tkinter para el funcionamiento de la misma. Se utilizó esta herramienta para hacer una mejor disposición de los botones y áreas de trabajo. También se pueden manipular algunas configuraciones como el tamaño, fuente y color del texto mostrado en la interfaz, además de tamaño de botones, al igual que el tamaño de ventana.

La comunicación con el ambiente ROS es una parte fundamental para el funcionamiento de la interfaz, así que ese fue el primer objetivo al desarrollar la GUI. Se logró de una manera sencilla pero funcional. El gran problema era que ROS y la GUI necesitan un ciclo independiente para poder funcionar, manejar varios hilos hacía de nuestro programa fuera inestable, así que se probó tener un solo ciclo para hacer funcionar ambas partes y funcionó de manera correcta. Es así como la interfaz continúa funcionando y de manera estable.

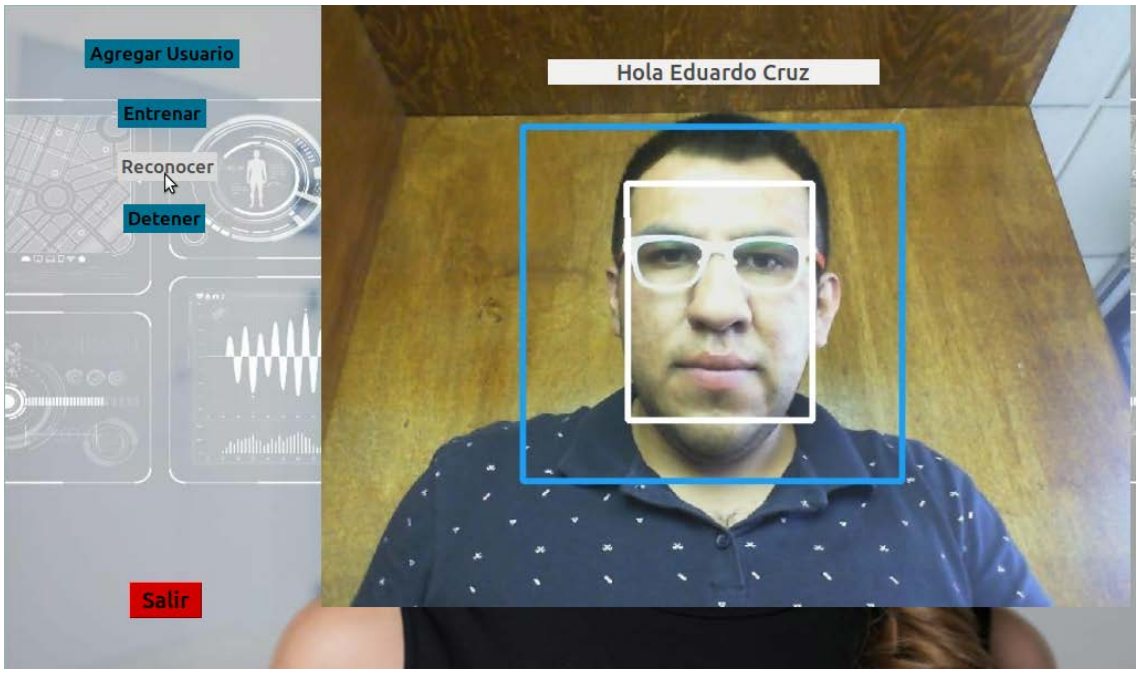


FIGURA 56 VISTA DE LA GUI EN LA FUNCIÓN RECONOCER

Para facilitar la obtención de resultados se desarrolló una versión de la interfaz en la cual se agregaron 3 botones, con los cuales el usuario o administrador indican si el reconocimiento fue correcto, incorrecto, o aparece como usuario no registrado. Una vez que se está en una rama que contenga esta configuración, simplemente se entra en modo reconocimiento y aparecerán los botones para ir agregando registros en los archivos de estadísticas.

FIGURA 57 VISTA DE LA GUI EN MODO ESTADÍSTICAS.

10. ANÁLISIS DE RESULTADOS

En esta etapa se presentan tres versiones del proyecto. Se realizaron poco más de 210 pruebas para cada una de las versiones, un total de 640 pruebas. Hay que recordar que, para reconocer a una persona, se toman diez imágenes, se clasifican las diez imágenes y se selecciona la clase con mayor índice de clasificación. De forma que, a pesar de realizar 640 pruebas, se procesaron 6400 imágenes.

10.1 Resultados PCA con 10 usuarios registrados.

	precision	recall	f1-score	support
Abraham Rodriguez	0.86	1.00	0.92	12
Angel Villalobos	1.00	0.50	0.67	18
Eduardo Angel	1.00	0.73	0.84	11
Eduardo Cruz	1.00	0.79	0.88	24
Hector Martinez	0.50	0.12	0.20	8
Luis Fortanel	0.92	1.00	0.96	11
Naham Tornes	0.00	0.00	0.00	0
Usuario No Registrado	0.83	0.87	0.85	127
Vanesa Barrueta	0.00	0.00	0.00	0
avg / total	0.86	0.81	0.82	211

FIGURA 58 REPORTE DE CLASIFICACIÓN PCA 10 USUARIOS

Previamente se mostraron los resultados obtenidos con LDA y PCA; sin embargo, en las pruebas finales se agregó la nueva clase denominada “Usuario no registrado”.

Para clasificar a los usuarios no registrados, se utilizó el método de clase “*predict_proba*”, del clasificador SVM. Se eligió un umbral de 0.4, de tal forma que si la probabilidad de pertenecer a una clase es mayor de 0.4, entonces la imagen es clasificada según la clase predicha. Si la probabilidad es menor a 0.4, entonces es clasificado como “Usuario no registrado”.

Precision y *recall,* son de gran importancia para evaluar el rendimiento del sistema.

La relevancia del *recall* de la clase “Usuario no registrado” es de gran importancia, pues este mide el desempeño frente a usuarios no registrados a los que se les otorgó el acceso.

En la matriz de confusión se puede notar que, de 127 intentos de usuarios no registrados, se les otorgó el acceso 17 veces.

Por otra parte, de 84 intentos de usuarios registrados, 23 veces se detectó como no registrados. Esto equivale al 72% de clasificación correcta para usuarios registrados.

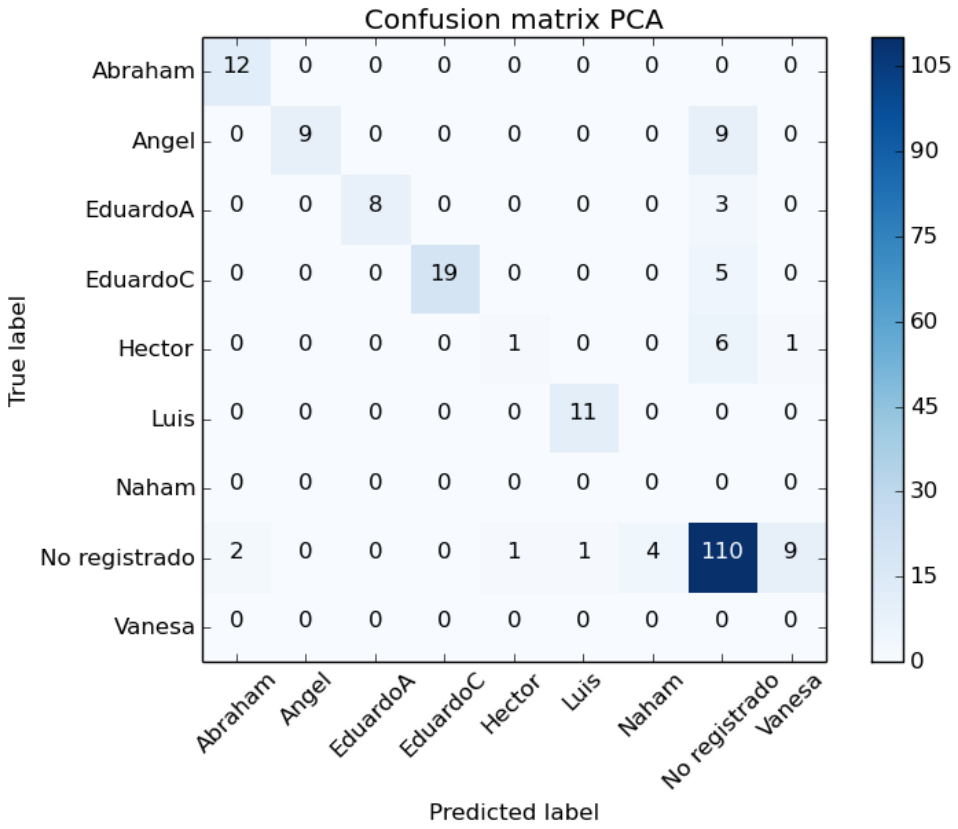


FIGURA 59 MATRIZ DE CONFUSIÓN PCA 10 PERSONAS.

10.2 Resultados LDA con 10 usuarios registrados.

	precision	recall	f1-score	support
Abraham Rodriguez	0.75	0.90	0.82	10
Angel Villalobos	1.00	0.85	0.92	13
Eduardo Angel	0.69	0.75	0.72	12
Eduardo Cruz	1.00	0.71	0.83	31
Hector Martinez	0.00	0.00	0.00	0
Luis Fortanel	1.00	0.85	0.92	13
Naham Tornes	0.00	0.00	0.00	0
Ricardo Elizalde	0.00	0.00	0.00	0
Usuario No Registrado	0.88	0.85	0.86	132
Vanesa Barrueta	0.00	0.00	0.00	0
avg / total	0.89	0.82	0.85	211

FIGURA 60 REPORTE DE CLASIFICACIÓN PCA 10 USUARIOS

En comparación con los resultados utilizando PCA, específicamente si observamos la clase “Usurario no registrado”, la precisión utilizando LDA es más alta, esto significa que los usuarios registrados fueron “confundidos” en un menor índice respecto a PCA.

Si notamos el *recall*, es el caso opuesto. El índice de *recall* es menor, esto significa que hubo un mayor número de casos en los que usuarios no registrados obtuvieron el acceso al sistema. Dicha información se puede notar en la matriz de confusión.

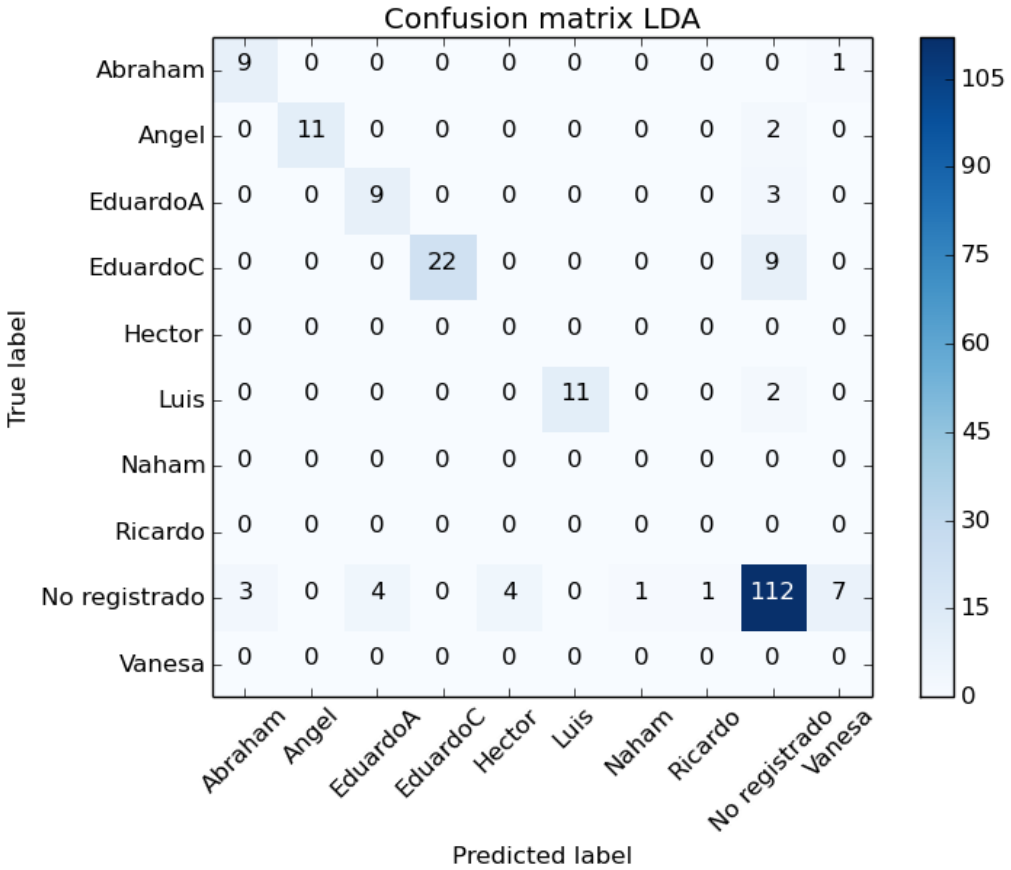


FIGURA 61 MATRIZ DE CONFUSIÓN LDA 10 PERSONAS.

También se puede notar que, de 132 intentos de usuarios no registrados, se les otorgó el acceso 20 veces.

Por otra parte, de 79 intentos de usuarios registrados, 17 veces se detectó como no registrados. Esto equivale al 86% de clasificación correcta para usuarios registrados.

La clasificación utilizando PCA como método de extracción de características tuvo un mejor desempeño contra los usuarios no registrados en la base de datos; sin embargo, la clasificación utilizando el método LDA es más eficiente para reconocer a los usuarios que sí están registrados.

10.3 Resultados PCA con 17 usuarios registrados.

Uno de los objetivos es almacenar e implementar el sistema para al menos 10 usuarios. A continuación, se presentan los resultados de 17 usuarios registrados en la base de datos. Sólo se registraron las estadísticas de 11 de los usuarios registrados.

	precision	recall	f1-score	support
Abraham Rodrioguez	0.79	0.92	0.85	12
Alejandro Pena	1.00	0.74	0.85	19
Andrea Ponce	1.00	0.43	0.60	14
Angel Villalobos	1.00	0.31	0.47	13
Eduardo Angel	1.00	0.44	0.62	18
Eduardo Cruz	0.94	0.77	0.85	22
Gerardo Fosado	1.00	0.68	0.81	19
Hector Martinez	0.00	0.00	0.00	8
Luis Forntanel	1.00	0.71	0.83	14
Octavio Montes	0.00	0.00	0.00	0
Pablo Torres	0.00	0.00	0.00	0
Usuario No Registrado	0.56	0.90	0.69	79
avg / total	0.79	0.71	0.70	218

FIGURA 62 REPORTE DE CLASIFICACIÓN PCA 17 USUARIOS

Nuevamente, si se observa específicamente la clase “Usuario no registro”, basta con notar la precisión para saber que el sistema clasificó incorrectamente en repetidas ocasiones a los usuarios que sí están registrados; sin embargo, mejoró el índice de usuarios no registrados clasificados de manera correcta.

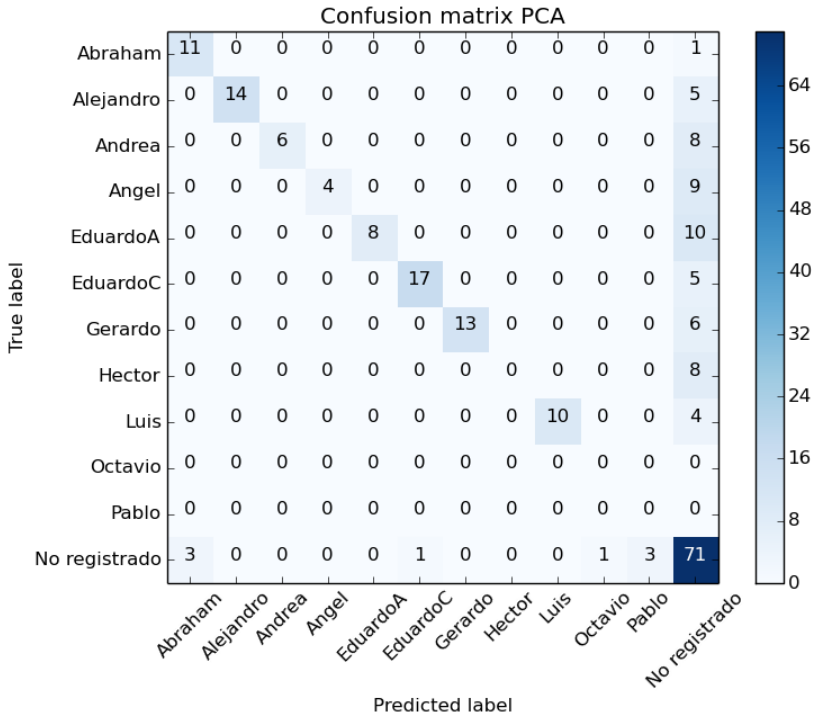


FIGURA 63 MATRIZ DE CONFUSIÓN PCA 17 USUARIOS.

11. GUÍA DE USO DE LA INTERFAZ DE USUARIO

Elementos de la interfaz de usuario.

Esta es la pantalla principal de la interfaz en ella se encuentran 5 botones en la parte izquierda que son; **Agregar usuario, Entrenar, Reconocer, Detener y Salir**. En la parte derecha se muestran diversos elementos según la funcionalidad que se esté ejecutando.

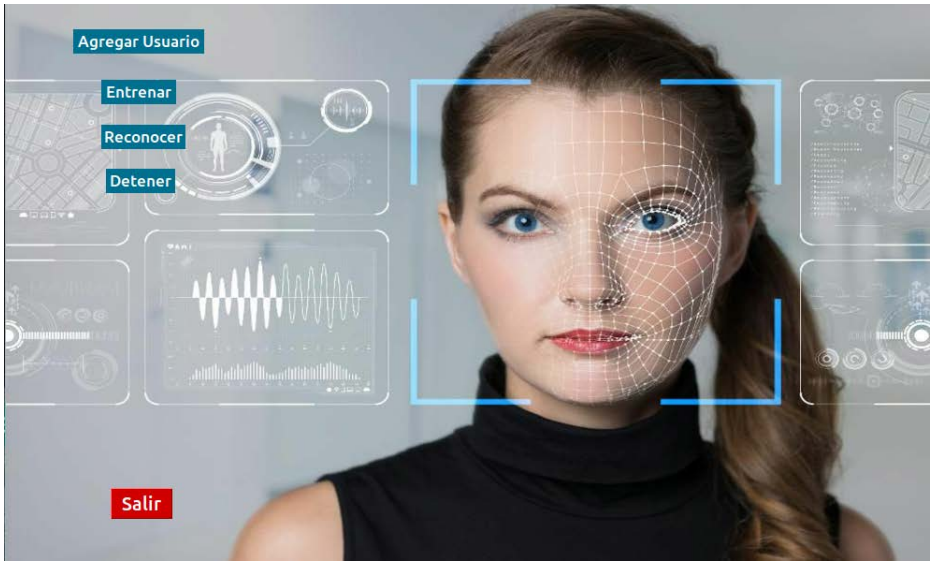


ILUSTRACIÓN 1 VISTA PRINCIPAL DE LA GUI

Agregar un usuario a la base de datos.

Para agregar un usuario a la base de datos, se deben seguir las siguientes instrucciones.

Presionar botón **Agregar Usuario**, aparecerá un cuadro de dialogo como el que se muestra en la imagen.

ILUSTRACIÓN 2 ESCRIBIR NOMBRE DEL NUEVO USUARIO (MODO AGREGAR USUARIO)

Escribir el nombre del nuevo usuario en el cuadro de dialogo, y presionar el botón **Crear Usuario**.

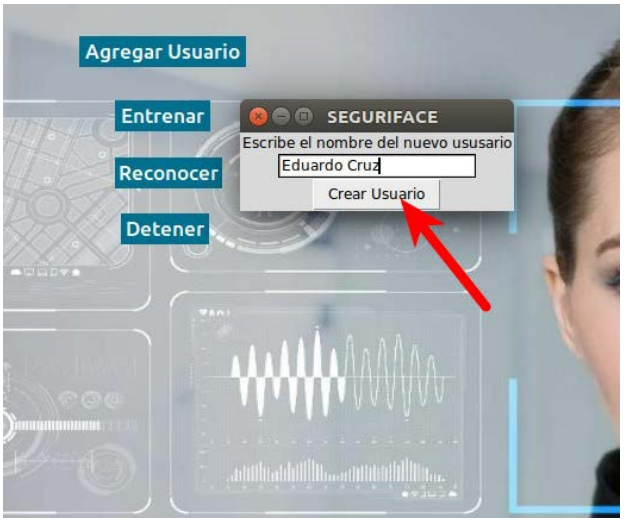


ILUSTRACIÓN 3 NOMBRE DEL NUEVO USUARIO, ANTES DE EMPEZAR EL GUARDADO DE IMÁGENES.

Se iniciará la cámara, el usuario debe colocar su rostro dentro del cuadro azul para que se empiecen a guardar imágenes en la base de datos. Este proceso tarda en promedio 25 segundos. No es necesario presionar el botón “detener”; al juntarse todas las imágenes el sistema detiene la captura de imágenes automáticamente.

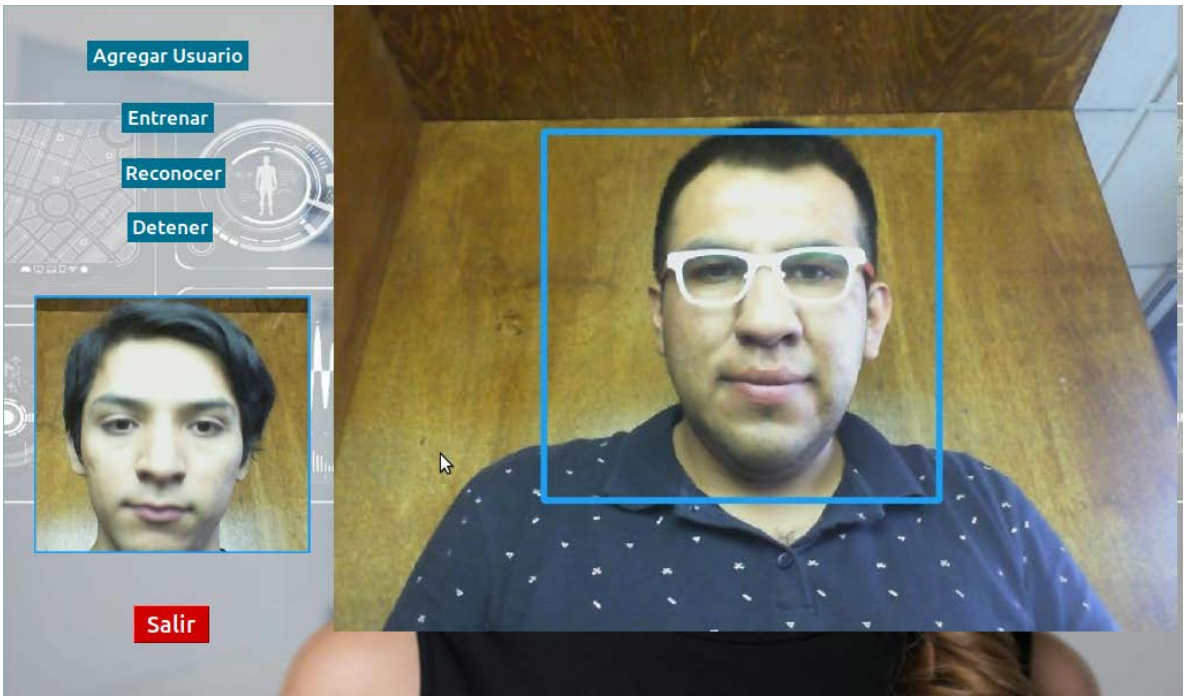


ILUSTRACIÓN 4 CAPTURA DE IMÁGENES DE LA BASE DE DATOS DEL NUEVO USUARIO.

Nota: no haga movimientos bruscos del rostro, no gire el rostro, y no haga gestos que comúnmente no hace, trate que su rostro se encuentre dentro del cuadro azul hasta que la imagen desaparezca.

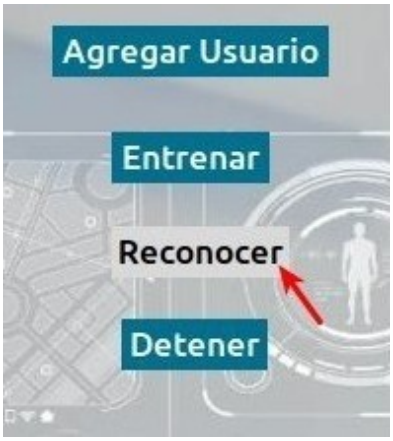
Entrenar el sistema.

Una vez que se han agregado nuevos usuarios al sistema es necesario crear un nuevo modelo, para poder ser reconocidos. Para ello, se debe presionar el botón **Entrenar** y esperar hasta que se muestre una ventana con varias imágenes de test.

ILUSTRACIÓN 5 MODO ENTRENAR.

Modo Reconocimiento y acceso.

Se debe presionar el botón **Reconocer**, esperar que la iluminación se estabilice, colocar su rostro dentro del sistema y parpadear hasta que el sistema reconozca su parpadeo. Una vez reconocido su rostro, su nombre aparecerá en la parte superior de la imagen. Si todo sale bien, oirá una chicharra, el cual es un indicador que puede pasar simplemente empujando la puerta.



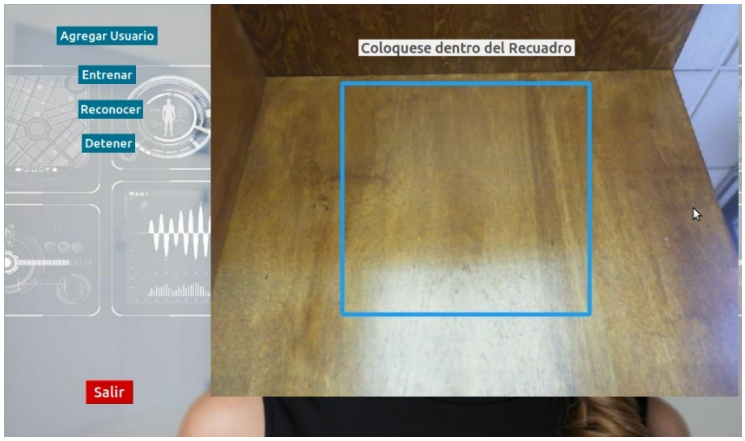


ILUSTRACIÓN 6 MODO RECONOCER, ESPERANDO QUE ALGUIEN COLOQUE SU ROSTRO DENTRO DEL RECUADRO AZUL

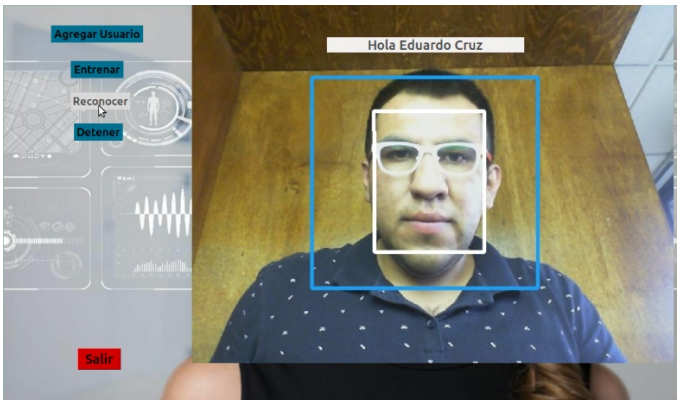


ILUSTRACIÓN 7 USUARIO RECONOCIDO EN MODO RECONOCER.

Detener Procesos.

Para salir del modo reconocimiento y agregar un usuario o volver a entrenar, se debe presionar el botón **Detener**. Para salir del modo agregar usuario por algún error, se debe presionar el botón detener y borrar el último registro en el archivo users.csv.

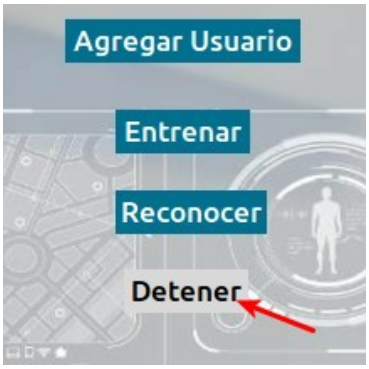


ILUSTRACIÓN 8 DETENIENDO PROCESOS.

Salir (Apagar).

Para detener todos los procesos y cerrar el sistema, solo se debe presionar el botón Salir.

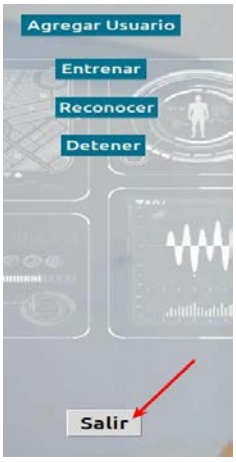


ILUSTRACIÓN 9 APAGANDO EL SISTEMA.

12. CONCLUSIONES

Este proyecto nos ha dejado con experiencia y aprendizaje en la implementación de herramientas enfocadas al reconocimiento facial y detección de rostro genuino. A continuación, se describen las conclusiones más importantes de este proyecto.

El punto más relevante fue el evaluar la seguridad del sistema de reconocimiento, se concluyó que agregando a más personas a la base de datos se reduce la efectividad del sistema, pasando de un promedio de 81% de *recall* con 10 usuarios registrados, a 71 % con 17 usuarios en la base de datos. Todavía se cumple con el objetivo propuesto en ambos casos, pero poniéndose en riesgo si se agregaran más usuarios a la base de datos.

Se trabajó con dos métodos de extracción de características, PCA y LDA los cuales se mencionan en el desarrollo de este proyecto. Se implementaron los dos métodos por separado, ya que el test realizado durante el entrenamiento de cada uno no arrojó resultados que diferenciaran determinantemente qué método era mejor para implementar. Al implementar ambas técnicas, se encontró que cada método tiene ciertas ventajas. PCA tiene un menor índice de reconocimiento, pero a la vez tiene menos falsos positivos, en contraste con LDA el cual tiene un mayor índice reconocimiento, pero mayor número de falsos positivos. Ninguno de los resultados antes mencionados permite concluir que un método es mejor a otro, así que esto se deja a consideración de las futuras aplicaciones que pueda tener el sistema.

La siguiente conclusión está relacionada con la verificación del rostro genuino en imágenes 2D. Al implementar un método híbrido se obtuvo un grado de seguridad satisfactorio. Se combinó un método de software basado en movimiento y uno por hardware, el primero consiste en un detector de parpadeos, el cual es necesario para que el sistema otorgue el acceso aun cuando se haya reconocido de manera correcta a una persona. El segundo consiste en un sensor de distancia que indica si existe un objeto plano como una fotografía frente al sensor. Ambos métodos establecen condiciones necesarias para controlar el acceso del usuario.

Se instaló una cabina para emular el entorno final en el que tendría lugar el control el acceso, se observó la influencia de la iluminación externa cambiante a lo largo del día y como el método de ecualización adaptativo en conjunto con la iluminación seleccionada permitieron contrarrestar defectos en las imágenes, un problema frecuente en sistemas de reconocimiento facial.

El proceso de control de acceso se realiza en aproximadamente 5 segundos, aunque esto puede variar en consecuencia del tiempo que tarda una persona en parpadear, esto aumenta el tiempo de en qué se completa el proceso ya que solo la etapa de extracción de características y clasificación se realiza en menos tiempo.

Por otra parte, se logró que el funcionamiento del sistema se controlara desde una interfaz gráfica, desde la cual se puede agregar usuarios nuevos, entrenar el sistema y reconocer, lo cual es importante ya que el usuario final no necesita tener conocimientos relacionados con el reconocimiento facial para hacer uso del sistema.

13. TRABAJOS FUTUROS

Durante la realización de este proyecto se encontraron ciertos aspectos que pueden ser tomados en cuenta para mejorar el rendimiento del sistema. Por ejemplo, aumentar la cantidad de usuarios registrados y sistemas simultáneos. A continuación, se explican ciertos trabajos futuros que se pueden tomar en cuenta para continuar con el desarrollo y mejoramiento de este sistema.

Gran parte de la obtención de buenos resultados en reconocimiento facial, depende de la creación de una base de datos de calidad. Se deben evitar gestos que naturalmente no se hacen, además de mantener una posición de la cara lo más vertical posible. Se sabe que no siempre se pueden controlar esos aspectos. La primera mejora al sistema consiste en la rotación de rostro, mediante la implementación de un detector de ojos, en el que se utilice su posición para contrarrestar el giro en el rostro, obteniendo una base de datos con imágenes más homogéneas en cuanto a alineación de rostro.

El segundo punto y más importante, desde nuestro punto de vista, se relaciona con la cantidad de usuarios registrados. Como se mencionó en los resultados, se obtuvieron mejores resultados con una base de datos de 10 personas que con una base de datos de 17 personas. Esta tendencia se debe al aumentar las clases de la base de datos, por lo tanto, no es conveniente usar nuestro sistema para aplicaciones donde se tengan que registrar más de 10 personas, esto para asegurar un buen funcionamiento. A continuación, se plantea una propuesta que no limitaría la cantidad de usuarios, excepto por la capacidad de memoria del sistema. Se propone una relación entre una tarjeta y el rostro de un usuario. Se plantea que un código de barras o una tarjeta RFID pueden funcionar como identificación del usuario, en conjunto con el reconocimiento facial requerido para otorgar el acceso. Con esto se sabría con quién comparar la persona que se quiere reconocer, haciendo el sistema más seguro y robusto al momento de otorgar el acceso.

Por último, y con los dos aspectos antes mencionados e implementados, se podría realizar una red de controles de accesos, en los que exista una sola base de datos. Gracias a la utilización de ROS, el sistema puede ser fácilmente escalable para controlar los accesos de todo un edificio. Esta sería la aplicación ideal del sistema, ya que se estarían explotando al máximo todas las herramientas utilizadas para el desarrollo de este proyecto.

Se ha dejado una buena base de este proyecto y se han utilizado herramientas que son de fácil interpretación para la continuación del proyecto. Cualquier persona interesada en él, puede clonar nuestro repositorio en GitHub y hacer sus propias modificaciones e implementaciones, por lo pronto quedamos más que satisfechos con el trabajo realizado y llenos de motivación para seguirnos retando en futuros proyectos.

14. REFERENCIAS

[1] R. G. Hernández, «Estudio de Técnicas de Reconocimiento Facial,» *Universitat Politecnica de Catalunya,* p. 86, 2010.

[2] AURORA, «Face Recongnition Biometric Control Reader,» [En línea]. Available: http://www.auroracs.co.uk/access-control/. [Último acceso: 10 02 2017].

[3] F. d. I. UNAM, «Biometría Informática. Clasificación de los sistemas biométricos,» [En línea]. Available: http://redyseguridad.fi-p.unam.mx/proyectos/biometria/clasificacionsistemas/clasificaciontecnologia.html. [Último acceso: 20 03 2017].

[4] F. Argentina, «Hoja técnica de equipos, Equipo Modelo Iface 302,» [En línea]. Available: http://www.fingertech.com.ar/equipos-autonomos/biometricos/Hoja-tecnica-equipos-biometricos/Hoja-tecnica-Iface-302.pdf. [Último acceso: 10 02 2017].

[5] F. I. Products, «Honor Serials Access Control and Time Attendance - Model F9010,» [En línea]. Available: http://www.faceidglobal.com/list-24-1.html. [Último acceso: 9 02 2017].

[6] K. Portugal, «Reconocimiento facial 3D para control de acceso y presencia en hospitales y farmaceuticas,» [En línea]. Available: http://www.kimaldi.com/aplicaciones/reconocimiento_facial/. [Último acceso: 10 02 2018].

[7] C. Villegas Quezada, «Reconocimiento de rostros utilizando análisis de componentes principales: limitaciones del algoritmo,» Universidad Iberoamericana, México, D.F., 2005.

[8] A. L. García Hernández, «Sistema de reconocimiento facial,» UPIITA IPN, México, D.F., 2004.

[9] A. A. Cortés Rosales, «Sistema de seguridad mediante el reconocimiento ocular,» UPIITA IPN, México, D.F., 2009.

[10] S. d. R. F. T. d. V. tridimensional, «Miguel Angel Vazquez Lopez,» *Centro de Investigaciones en Optica,* 2014.

[11] M. Á. V. López, Sistema de Reconocimiento Facial Mediante Técnicas de Visión Tridimensional, Guanajuato: Centro de Investigciones en Optica A.C., 2014.

[12] V. Vapnik, «Support-vector networks,» de *Machine Learning*, 1995.

[13] M. Grgic, "FACE RECOGNITION," 18 enero 2017. [Online]. Available: http://www.face-rec.org/algorithms/. [Accessed 2017 marzo 20].

[14] F. F. Li, «YouTube,» TED TALKS, 23 marzo 2015. [En línea]. Available: https://www.youtube.com/watch?v=40riCqvRoMs. [Último acceso: 13 marzo 2017].

[15] «RASPBERRY PI 3 MODEL B,» Fundacion Raspberry Pi, [En línea]. Available: https://www.raspberrypi.org/products/raspberry-pi-3-model-b/. [Último acceso: 11 11 2017].

[16] I. G. Marcos, Tecnicas y Algoritmos de Vision Artificial, España: Universidad de La Rioja, 2016.

[17] S. M. M. H. Sajida Parveen, «Face anti-spoofing methods,» *Current Science,* vol. 108, nº 8, 2015.

[18] L. S. Z. W. y. S. L. Pan, «Eyeblink-based anti-spoofing in face recognition from a generic webcamera,» *Proceedings of the IEEE 11th International Conference on Computer Vision ,* nº 7, pp. 1-8.

[19] Y. W. T. T. a. A. K. J. J. Li, «Live face detection based on the analysis of Fourier spectra,» *Proceedings of the Biometric Technology for Human Identification,* vol. 5404, p. 296–303, 2004.

[20] C. j. p. herrera, «ADABOOST – FILTRO DE BASE HAAR,» 27 04 2017. [En línea]. Available: https://carlosjuliopardoherrera.wordpress.com/2017/04/27/adaboost-filtro-de-base-haar/.

[21] OpenCV, «OpenCV,» 10 10 2017. [En línea]. Available: https://docs.opencv.org/master/d7/d8b/tutorial_py_face_detection.html.

[22] R. C. G. a. R. E. Woods, Digital Image Processing, Tercera edición, 2008.

[23] M. Kirby and L. Sirovich, «Application of the KL Procedure for the Characterization of Human Faces,» *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 12, nº 1, pp. 130-108, 1990.

[24] M. K. L. Sirovich, «Low-Dimensional Procedure for Characterization,» *J. Optical Soc. Am.,* vol. 4, pp. 519-524, 1987.

[25] M. Turk and A. Pentland, «Eigenfaces for Recognition,» *J. Cognitive Neuroscience,* vol. 3, nº 1, pp. 71-86, 1991.

[26] M. Grudin, «On Internal Representations in Face Recognition Systems,”,» *Pattern Recognition,* vol. 33, nº 7, pp. 1161-1177,, 2000.

[27] Y. Y. L. Zhao, «Theoretical Analysis of Illumination in PCA-Based Vision Systems,» *Pattern Recognition,* vol. 32, nº 4, pp. 547-564, 1999.

[28] O. A. Reference, «Face Recognition with OpenCV,» OpenCV 2.4.13.4 documentation , [En línea]. Available: https://docs.opencv.org/2.4/modules/contrib/doc/facerec/facerec_tutorial.html. [Último acceso: 23 10 2017].

[29] D. P. A. Collantes, «Reconocimiento facial basado en el analisis de componentes principales,» *Facultad de Ciencias. Universidad Nacional de Trujillo,* vol. 1, nº 1, pp. 28-30, 2013.

[30] D. Z. A. F. Jian Yang, «Two-dimensional PCA: a new approach to appearance-based face representation and recognition,» *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE,,* 2014.

[31] M.Welling, «Fisher Linear Discriminant Analysis",» *Department of Computer Science, University of Toronto.*

[32] R. J. Q. L. Jieping Ye, «Two-Dimensional Linear Discriminant Analysis,» *NIPS'04.*

[33] X. H. &. P. Niyogi., «Locality Preserving Projections,» *NIPS'03,* vol. 16.

[34] J. N. T. T. G. X. a. S. Y. Q. Bai, «Is physics-based liveness detection truly possible with a single image?,» *the IEEE International Symposium. Circuits and Systems,* p. 3425–3428., 2010.

[35] K. F. H. a. B. J. Kollreider, «Non-intrusive liveness detection by face images.,» *Image Vision Comput.,* nº 27, p. 233–244., 2009.

[36] W. R. R. A. Schwartz, «Face spoofing detection through partial least squares and low-level descriptors.,» *International Joint Conference on Biometrics (IJCB),,* 2011.

[37] T. F. A. A. D. M. J. M. a. S. M. Pereira, «LBP-TOP based countermeasure against face spoofing attacks,» *ACCV 2012 Workshop,* vol. 1, pp. 121-132, 2012.

[38] I. A. A. a. M. S. Chingovska, «On the effectiveness of local binary patterns in face anti-spoofing.,» *International Conference on Biometric Special Interest Group,* 2012.

[39] N. a. M. S. Erdogmus, «Spoofing in 2D face recognition with 3D masks and anti spoofing with Kinect.,» *International Conference on Biometrics:,* 2013.

[40] K. F. H. F. M. a. B. J. Kollreider, «Real time face detection and motion analysis with application in liveness assesment,» *Trans. Infor. Forensics and Security, IEEE,* vol. 2, p. 548, 2007.

[41] J. S. y. J.Vandewalle, «Least Squares Support Vector Machine Classifiers,» *Neural Processing Letters,* nº 9, 1999.

[42] Scikit-learn, «scikit-learn,» 14 11 2017. [En línea]. Available: http://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html.

[43] element14, «Element14,» 20 08 2017. [En línea]. Available: https://www.element14.com/community/community/raspberry-pi?src=raspberrypi. [Último acceso: 20 08 2017].

[44] raspberryparatorpes, «raspberryparatorpes,» 20 08 2017. [En línea]. Available: https://raspberryparatorpes.net/rivales/rivales-raspberry-pi-odroid-xu4/. [Último acceso: 20 08 2017].

[45] slant, «slant,» 22 08 2017. [En línea]. Available: https://www.slant.co/versus/5727/15521/~odroid-xu4_vs_raspberry-pi-3-model-b. [Último acceso: 22 08 2017].

[46] Microsoft, «Microsoft Accessories,» 28 08 2017. [En línea]. Available: https://www.microsoft.com/accessories/es-es/products/webcams/lifecam-cinema/h5d-00002. [Último acceso: 29 08 2017].

[47] V. B. H. Hill, «EfFects of Lighting on Matching Facial Surfaces,» *Journal of Experimental,* vol. 22, pp. 986-1004, 1996.

[48] Iluminet, «Norma Mexicana para la correcta iluminación de los sitios de trabajo,» La NOM 025-STPS-2008 rige los lineamientos básicos para la iluminación de sitios de trabajo., [En línea]. Available: http://www.iluminet.com/nom-025-iluminacion-trabajo/. [Último acceso: 10 11 2017].

[49] M. A. B. I. Rahimeh Rouhi, «A review on feature extraction techniques in face recognition,» *Signal & Image procesing: An international Journal,* vol. 3, nº 6, p. 15, 2012.

[50] J. A. U. Abad, «Universidad Politecnica de Madrid,» 09 2014. [En línea]. Available: http://oa.upm.es/33688/1/PFC_juan_alfonso_urtiaga_abad.pdf. [Último acceso: 5 03 2018].

[51] A. M. H. Omar Faruqe, «Face Recognition Using PCA and SVM,» de *2009 3rd International Conference on Anti-counterfeiting, Security, and Identification in Communication*, IEEE, 2009.

[52] D. L. Salgado, «Visión Artificial:Fundamentos y Aplicaciones,» Universidad Politécnica de Madrid, Madrid, España, 2007.