

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01- 27

**COMPARAR TRES MÉTODOS CONVENCIONALES DE
RECONOCIMIENTO FACIAL Y CONFRONTAR EL DE MEJOR
RESULTADO FRENTE A COMPUTACIÓN COGNITIVA.**

Alexis Ospina Herrera

Mateo Pulido Morales

Ingeniería en Sistema

Director(es) del trabajo de grado

INSTITUTO TECNOLÓGICO METROPOLITANO

18 de Septiembre de 2017

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

RESUMEN

El proyecto consistirá en una comparación y análisis de 3 técnicas diferentes de reconocimiento y detección de personas en video, estos métodos consistirán en el uso de un clasificador, cualquiera de las metodologías del proyecto: el uso de máquinas de soporte de vectores, en cascada clasificadores y el uso de detección de puntos. Esto permitirá obtener información sobre la precisión y complejidad de cada técnica o metodología, y también puede obtener información sobre el entrenamiento cuando se ingresa a una nueva persona.

Para la ejecución de este proyecto se debe seguir un flujo muy estricto ya que cada paso que se realiza en el proyecto dependerá del paso anterior. Como una idea central del proyecto se desarrollará una comparación de las tres metodologías mencionadas anteriormente para encontrar el que tiene la mayor precisión y el mayor porcentaje de visitas. Inicialmente se construirá una base de datos de caras que serviría de insumo para establecer un número de imágenes por persona y el número de ellas que se utilizarán, esto dependerá de la metodología que se va a ejecutar, es decir, cada metodología se utilizará como introducir una cantidad diferente de imágenes para su propia implementación. El siguiente paso es la implementación de los diferentes algoritmos que corresponden a las diferentes metodologías con el apoyo de diferentes bibliotecas.

Con la implementación de cada una de las metodologías se podría realizar un análisis de los resultados obtenidos que permitan cuantificar su precisión o porcentaje de éxito, tiempo de ejecución, cantidad de imágenes (tablas) utilizadas y tiempo de reciclaje. Entonces será posible determinar qué metodología tiene mejores resultados que los demás en términos de porcentaje de identificación y reconocer las limitaciones de cada metodología implementada. Finalmente, habiendo identificado la mejor metodología con mejores resultados se hará una comparación con la implementación que se desarrolla a partir de la

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

computación cognitiva. Hoy en día, en la era de los sistemas cognitivos, los sistemas pueden aprender de sus experiencias, encontrar correlaciones, crear hipótesis y recordar los resultados y aprender de ellos un valor añadido al proyecto.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

RECONOCIMIENTOS

Este proyecto fue construido por la formación recibida durante toda la carrera tanto en la tecnología e ingeniería del ITM. Los diferentes profesores que participaron en nuestra formación desde el primer semestre aportaron su granito de arena para adquirir cada uno de los conocimientos que hoy en día nos transforma en profesionales.

Durante el proyecto agradecer especialmente a Henry Mauricio Vásquez Carvajal, quien fue la persona que nos dio el norte durante todo el proyecto, ante su apoyo y guiarnos durante cada etapa del proyecto con su experiencia y conocimiento; nunca recibimos comentarios de desánimo, si no de completo apoyo para sacar el proyecto adelante siendo certero en cada punto o idea que nos quería transmitir para ir firmes durante cada una de las fases del proyecto.

Además, queremos agradecer principalmente a nuestras familias por ser el pilar fundamental en todo lo que somos, en toda nuestra educación, tanto académica, como de la vida, por su incondicional apoyo perfectamente mantenido a través del tiempo.

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

ACRÓNIMOS

EIGEN: reconocimiento de caras por método Eigenfaces.

LBPH: reconocimiento de caras por método Local Binary Patterns Histograms.

FISHER: reconocimiento de caras por método fisherfaces.

SVM: Support Vector machines.

SURF: Detector de puntos Speeded-up Robust Features.

KEYPOINTS: puntos importantes o puntos claves.

XYZ: modelo de color xyz.

HLS: modelo de color matiz, saturación, luminosidad.

GFTTDETECTOR: función de apoyo de opencv Good Features To Track.

JSON: JavaScript Object Notation.

RBF: función de base radial o Radial basis function.

TRHESHOLD: umbral.

INTER: intersección de histograma o Histogram intersection kernel.

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

TABLA DE CONTENIDO

1.	INTRODUCCIÓN	8
1.1	PLANTEAMIENTO DEL PROBLEMA	9
1.2	OBJETIVOS	9
1.2.1	GENERAL	9
1.2.2	ESPECÍFICOS	9
2.	MARCO TEÓRICO	11
	ESTADO DEL ARTE	11
2.1	BASE DE DATOS DE ROSTROS	13
2.2	OPENCV	13
2.2.1	CLASIFICACIÓN DE VECTORES DE SOPORTES LINEALES Y NO LINEALES	15
2.3	MACHINE LEARNING (APRENDIZAJE DE MÁQUINA)	16
2.4	MÁQUINAS DE APRENDIZAJE	16
2.5	PROCESAMIENTO MORFOLÓGICO	18
2.5.1	DILATACIÓN	18
2.5.2	EROSIÓN	19
2.5.3	APERTURA Y CIERRE	20
2.5.4	FILTROS DETECTORES DE BORDES	22
2.6	ECUALIZACIÓN DE HISTOGRAMAS	23
2.7	MÉTODO RECONOCIMIENTO FACIAL EN CASCADA	24
2.7.1	DETECCIÓN DE ROSTRO	24
2.7.2	RECONOCIMIENTO FACIAL	26
2.7.3	EIGENFACES Y FISHERFACES	27
2.7.4	HISTOGRAMAS DE PATRONES BINARIOS LOCALES	29
2.8	DETECCIÓN FACIAL CON PUNTOS	30
2.8.1	SURF	31
2.8.2	HOMOGRFÍAS	31
2.8.3	DETECTORES DE PUNTOS DE INTERÉS	33
2.8.4	DESCRIPCIÓN DE PUNTOS DE INTERÉS	34
2.9	COMPUTACIÓN COGNITIVA	34
3.	METODOLOGÍA	36
3.1	ESQUEMA GENERAL DE LAS ETAPAS Y BASE DE DATOS	36
3.1.1	EXTRACCIÓN DE IMÁGENES:	37

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

3.2	METODOLOGIAS	37
3.2.1	MÁQUINAS DE VECTORES DE SOPORTE O SVM:.....	37
3.2.2	MÉTODO CASCADA:.....	43
3.2.3	MÉTODO DETECCIÓN POR PUNTOS (Homografía):	44
3.2.4	METODO DE DETECCION POR PUNTOS (sin homografía):	47
3.3	DIFERENTES TRATAMIENTOS DE LA IMAGEN.....	48
3.3.1	MÉTODOS DE COMBINACIONES.....	48
3.3.1.1	COMBINAR 1:	48
3.3.1.2	COMBINAR 2:	48
3.3.1.3	MORFOLOGÍA:.....	49
3.3.1.4	COLOR:	49
3.3.1.5	ECUALIZACIÓN:	49
3.3.2	DIAGRAMAS FLUJOS DE COMBINACIONES	50
3.3.2.1	COMBINAR 1:	50
3.3.2.2	COMBINAR 2:	51
3.3.2.3	MORFOLOGÍA:.....	51
3.3.2.4	COLOR:	52
3.3.2.5	ECUALIZACIÓN:	52
3.4	RECONOCIMIENTO FACIAL CON COMPUTACIÓN COGNITIVA	52
4.	RESULTADOS Y DISCUSIÓN.....	55
4.1	MÁQUINAS DE VECTORES DE SOPORTE (SVM).....	55
4.2	CASCADA.....	56
4.3	DETECCIÓN POR PUNTOS	59
4.4	RESUMEN RESULTADOS DE COMPARACIÓN.....	60
4.5	COMPARACIÓN DE TIEMPOS.....	62
4.5.1	REENTRENAMIENTO – ENTRENAMIENTO	63
4.6	RECONOCIMIENTO FACIAL COGNITIVO	65
4.6.1	COMPUTACIÓN COGNITIVA vs METODOLOGÍA CON MEJOR RESULTADO.....	65
5.	CONCLUSIONES, RECOMENDACIONES Y TRABAJO FUTURO	68
	CONCLUSIONES.....	68
	RECOMENDACIONES	69
	TRABAJOS A FUTUROS.....	70
	REFERENCIAS	72

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

1. INTRODUCCIÓN

El rostro humano es un objeto dinámico que tiene un alto grado de variabilidad en su aspecto que hace que su detección sea un problema difícil de tratar en la visión por computador. Inicialmente, el problema de detección de rostros en los sistemas de reconocimiento no recibió la atención necesaria y se basó en el hecho de que el rostro ya había sido detectado, solo en los años ochenta los primeros algoritmos, basados en técnicas heurísticas y antropométricas, cuando el desarrollo de algoritmos de detección de rostros comenzó a crecer, se propone una gran variedad de técnicas, desde algoritmos básicos de detección de bordes hasta algoritmos compuestos de alto nivel utilizando métodos avanzados de reconocimiento de patrones.

Aunque para nosotros los seres humanos realizar el reconocimiento facial es relativamente simple, el problema es muy difícil de resolver desde un punto de vista computacional. Todo cambio mínimo de luz ambiente, expresión de la cara, posición, inclinación, cabello, etc. Supone una imagen totalmente diferente. Naturalmente, siempre es posible restringir las condiciones de funcionamiento y obligar al individuo a reconocer que está en una habitación con una cierta iluminación o colocada en una cierta posición con respecto a la cámara.

Este proyecto se basará en una comparación de tres métodos, estos métodos se pueden resumir en el siguiente flujo básico:

1. Detección de la cara en una imagen.
2. Extracción de la zona de la imagen que contiene la cara: p. ej. con un rectángulo o una elipse.
3. Normalización.
4. Reconocimiento.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

La detección se realizará a través de una webcam que servirá como entrada para la aplicación para iniciar el flujo de detección de caras y clasificación de ésta. La normalización consiste en modificar o homologar la imagen de la cara a reconocer. Es decir, la imagen tendrá un proceso de transformación antes de generar una clasificación. En la normalización se puede contar con la imagen que tiene un tamaño común, intensidades en un cierto rango. (Este proceso es necesario para generar un equilibrio entre las imágenes que desea evaluar con las que generaron el entrenamiento). En otras palabras, la normalización es lo que busca facilitar el proceso posterior de reconocimiento.

Finalmente, el reconocimiento consiste, para cualquiera de los tres métodos que se pueden ver en el proyecto, comparar la imagen de entrada con las imágenes que se almacenan en la base de datos.

1.1 PLANTEAMIENTO DEL PROBLEMA

El problema a resolver consiste principalmente en la detección e identificación de caras de vídeo. Para ello realizaremos una comparación y análisis de los resultados entre las tres técnicas que realizan esta tarea; las técnicas que se compararán son: el método en cascada, el uso de support vector machines y el método de detección de puntos. Con el fin de poder mostrar cuál de ellos genera un mejor resultado.

1.2 OBJETIVOS

1.2.1 GENERAL

Realizar un análisis de datos con tres metodologías diferentes, en las que podemos realizar una comparación que nos permita ver cuál de ellas nos da el mejor resultado y el más eficiente para hacer una comparación con la computación cognitiva.

1.2.2 ESPECÍFICOS

- Comparar el tiempo de clasificación, la medida de eficiencia algorítmica, el número de falsos positivos, el tiempo de entrenamiento de las tres metodologías convencionales.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

- Apreciar el tiempo de respuesta para observar cuánto tiempo dura el entrenamiento para realizar una clasificación correcta.
- Identificar el tiempo de reentrenamiento de las metodologías a las que se pueden hacer, lo que nos permitirá obtener información sobre el desempeño del algoritmo.
- Implementar la computación cognitiva y realizar una comparación del método con mejor resultado con la computación cognitiva.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

2. MARCO TEÓRICO

ESTADO DEL ARTE

Para la detección y clasificación de caras se utilizarán tres metodologías que se han implementado desde hace algún tiempo, el primer método consiste en el uso de máquinas vectoriales de apoyo que según Monadjemi el uso de SVM consiste principalmente en tener una base de datos con una cantidad de unas 150 imágenes, ya que estas máquinas necesitan de una gran cantidad de variantes, ángulos, posiciones, para poder ser entrenadas y posteriormente detectar la cara del pueblo (Monadjemi, 2011). Monadjemi también aclara que al utilizar esta metodología debe tener en cuenta el fondo y la iluminación donde se encuentran las caras de las personas, ya que dependiendo de que tendría que transformar o normalizar las imágenes. Según Zhang, estas normalizaciones consisten principalmente en generar histogramas a las imágenes, lo que permite que las máquinas vectoriales de soporte reconozcan las imágenes más fácilmente, basándose en estadísticas de dichos histogramas (Zhang, 2008); Este método tiene una tasa de éxito de entre el 80-90% de precisión, pero carece de velocidad en la ejecución.

La segunda metodología es el uso de un clasificador en cascada según Zhang al utilizar esta metodología, es posible implementar algunos módulos que permiten la extracción de características que pueden ser el color de la piel, la simetría en las diferentes partes del rostro como ojos, boca, nariz, entre otros (Zhang, 2008 , pág. 25). Para este clasificador también es necesario tener una base de datos amplia, ya que como el primer método necesita variantes para reconocer las diferentes posiciones y ángulos en los que se encuentra la cara, también pasa por una serie de transformaciones y normalizaciones con el uso de histogramas; pero con la diferencia de que este método utiliza un discretizador basado en si cumple o no ciertas condiciones. Según Cevikalp, Evangelos y Aliradi obtienen mejores resultados al utilizar el clasificador con características que incluyen ojos, boca y nariz. Especialmente si se puede separar o dividir en un conjunto que implica ciertas

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

particularidades que permiten la discretización de dichas partes. Este método tiene un porcentaje de éxito similar al primer método entre 75-90% pero carece de facilidad de entrenamiento cuando se requiere que una nueva persona sea clasificada (Aliradi, 2013 ; Cevikalp, Triggs, & Franc, 2013; Sariyanidi, 2012).

La tercera metodología consistirá en un algoritmo que realizará la detección por puntos, lo que, a diferencia de los otros 2 métodos mencionados anteriormente, esto no requiere una base de datos de gran número de imágenes y sólo con 2 imágenes puede reconocer a la persona; de modo que las variantes de ángulos y posiciones no necesitan ser muchas. Según Patil, este método consiste principalmente en extraer puntos de la cara que involucran diferentes partes, tales como ojos, boca, nariz, cejas, entre otros que permiten obtener una discretización precisa y un conjunto de características que varían de persona a persona basada en la distancia entre estos puntos (Patil, Sahula, & Mandal, 2011). Según Jian-Zheng para mejorar los resultados que se pueden obtener con este método se requiere normalizar las imágenes y pasarlas a escala de grises, dando así una mayor obtención de puntos, lo que permite dividir las más fácilmente en diferentes regiones (Jian-zheng, 2011). Esto fue verificado por Panning, que dividió la cara en diferentes regiones y aumentó el porcentaje de éxito del método hasta el 80% en más de 700 imágenes en diferentes posiciones y ángulos (Panning, Al-Hamadi, & Michaelis, 2012). Finalmente, este método tiene una tasa de éxito de hasta un máximo del 95% y corrige los problemas de ejecución del primer método y entrenamiento del segundo método.

Después de la implementación y análisis de datos de las metodologías, se procederá a la implementación del reconocimiento facial con computación cognitiva con el fin de comparar los resultados utilizando una metodología convencional con sistemas cognitivos. Según A. Hussain, la computación cognitiva está relacionada con la inteligencia artificial, de manera que permite la creación de un sistema autónomo con capacidades cognitivas, que están involucradas en la selección de diferentes acciones, este nuevo sistema permitiría un marco de trabajo que implica múltiples módulos que incluyen varios aspectos del lenguaje y la visión, permitiendo así la creación de agentes autónomos (Hussain, 2013, pág. 5).

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

2.1 BASE DE DATOS DE ROSTROS

La base de datos actúa como un repositorio de información biométrica. Durante el proceso de almacenamiento, el conjunto de características extraídas de la muestra biométrica sin procesar (es decir, la plantilla) se almacena en la base de datos junto con alguna información individual (como nombre, número de identificación personal, dirección, entre otros). La captura de datos durante el proceso de registro puede o no ser controlada por un ser humano dependiendo de la aplicación final de esto.

Por ejemplo, un usuario que intente crear acceso a un área restringida para un nuevo individuo usando su huella digital tendrá que presentar sus credenciales no biométricas para probar que necesita acceso y luego introducir sus datos biométricos, todo en presencia de un recinto. En cambio, un usuario que genera una cuenta en una computadora con acceso a través del reconocimiento facial, genera una serie de fotografías como una base de datos sin ninguna supervisión.

2.2 OPENCV

OpenCV (Open Source Computer Vision) es una biblioteca de visión artificial de código abierto, escrita en lenguaje C y C++. Fue diseñado para lograr una alta eficiencia computacional centrándose en aplicaciones en tiempo real. La biblioteca tiene más de 500 funciones que abarcan muchas áreas de visión por computadora, incluyendo fabricación, inspección de productos, imágenes médicas, seguridad, interfaz de usuario, calibración de cámara, visión estéreo y robótica, entre otros. También incluye una biblioteca de aprendizaje de máquinas de propósito general. OpenCV surgió como una iniciativa de investigación de Intel para avanzar en aplicaciones de uso intensivo de la CPU.

Al principio, los objetivos de OpenCV fueron, entre otros, avanzar en la investigación de la visión por computador ofreciendo código abierto y optimizado para la infraestructura de visión básica; difundir el conocimiento de la visión por computadora mediante la entrega de una infraestructura común en la que los desarrolladores podrían trabajar, facilitando la lectura y la escritura; finalmente, buscamos avanzar aplicaciones comerciales basadas en la

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

visión por computador, haciendo que el código sea portátil, con un rendimiento optimizado y sin costo alguno.

OpenCV se estructura en cinco componentes principales, cuatro de los cuales se muestran y describen en la figura 1.

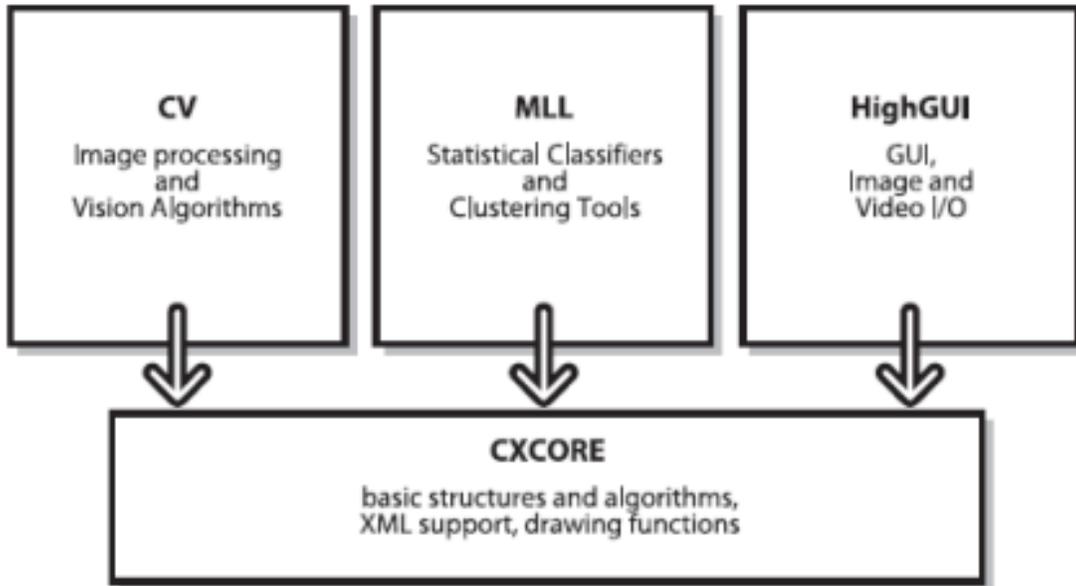


Figura 1. Estructura básica de OpenCV.

CV contiene funciones para el procesamiento básico de imágenes y los algoritmos de visión computarizada de más alto nivel; ML es la biblioteca de aprendizaje automático, que incluye varios clasificadores estadísticos y herramientas de agrupamiento; HighGUI contiene rutinas de entrada / salida y funciones para almacenar y cargar vídeos e imágenes y CXCore contiene las estructuras de datos y contenidos básicos. Otro componente es el CvAux, que contiene áreas extinguidas (Reconocimiento de Fallos) y algoritmos experimentales (Segmentación de Antecedentes y Primer Plano) (Bradski & Kaehler, 2008).

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

2.2.1 CLASIFICACIÓN DE VECTORES DE SOPORTES LINEALES Y NO LINEALES

Las máquinas vectoriales de soporte trabajan en dos tipos de casos, separables y no separables, en casos separables pueden usarse clasificadores lineales como figura 2, en el caso de núcleos no separables.

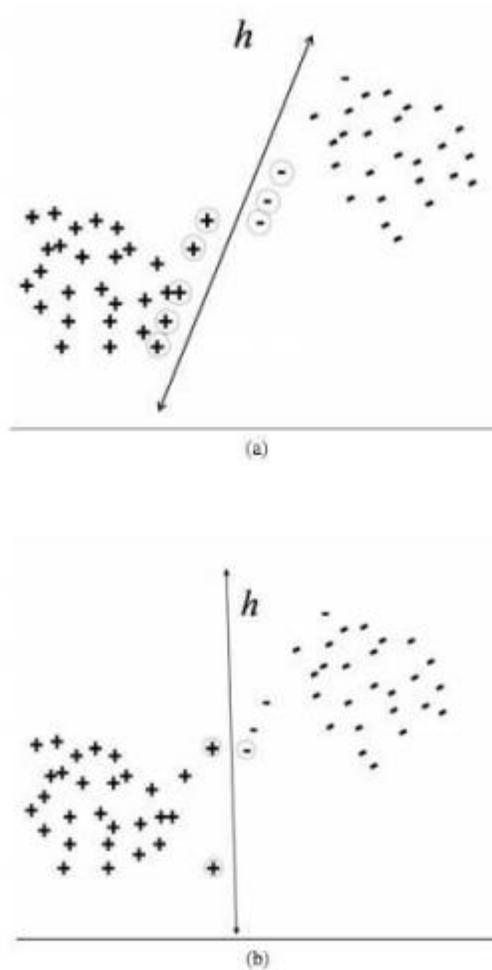


Figura 2: Ejemplos de hiperplano, a) el hiperplano que se aprendió por las SVM, los círculos corresponden a vectores de soporte, b) el aprendizaje de las SVM con un punto puede cambiar.

La separación óptima debe ser esencial, ya que permite elegir correctamente los vectores de soporte y hacer una buena clasificación de los datos del espacio de características. (Cohen, Sebe, Garg, Chen, & Huang, 2003)

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

2.3 MACHINE LEARNING (APRENDIZAJE DE MÁQUINA)

Estudia cómo la computadora puede simular o realizar el comportamiento humano, busca establecer modelos computacionales o entender modelos según el estudio de los mecanismos humanos a través de la psicología, la ciencia cognitiva y el estudio metodológico.

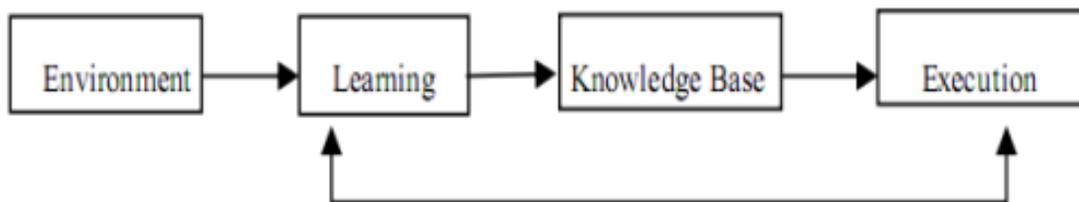


Figura 3: Estructura básica de un sistema de aprendizaje.

El entorno provee cierta información a la sección de aprendizaje del sistema y esta revisa la librería de conocimiento para usar dicha información y mejorar la parte de ejecución. El factor más importante para el diseño del sistema de aprendizaje es seleccionar la información que se proveerá al sistema y si dicha información es de calidad, también se tendrá en cuenta la creación de reglas correctamente constituidas que pueden hacer que el sistema sea eficiente, no obstante, unas reglas incorrectas pueden modificar o borrar información de la base de conocimiento. (Xue & Zhu, 2009)

El segundo factor importante que puede afectar el diseño del sistema es la base de conocimiento, el conocimiento puede ser expresado en varias formas, instancias, vectores de características, producción de reglas de reconociendo y reglas semánticas. (Xue & Zhu, 2009)

2.4 MÁQUINAS DE APRENDIZAJE

El aprendizaje de maquina se puede dividir en dos tipos: supervisado y sin supervisión.

El primero de estos tipos es el supervisado que está empleado en la ciencia porque permite el aprendizaje de patrones o características que no son evidentes para el ser humano. Este

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

tipo de máquina de aprendizaje es ideal para tareas de clasificación. La forma en que se ejecuta este tipo de procedimientos consiste en un paso de entrenamiento en el que se presenta el algoritmo de datos ya analizados para que aprenda las características de cada clase de partícula.

Una vez entrenado suele haber una fase de prueba en el que se le vuelve a dar datos conocidos, pero se presentan sin solución de forma que el ordenador tiene que aplicar todo lo aprendido mediante máquinas de aprendizaje para poder descubrir qué tipo de partícula se produjo. Si supera esta fase con un alto porcentaje de acierto el algoritmo está ya preparado para enfrentarse a nuevos datos y analizarlos mejor.

Por otro lado, se encuentra el aprendizaje no supervisado en el que no existe una fase de entrenamiento. En este caso el algoritmo suele tener unos criterios de calidad de la respuesta y aprende cuál es la mejor manera de llegar al resultado óptimo lo más rápido posible. Es decir, el no supervisado solo hay muestras de entrada, pero no hay de salida durante el aprendizaje. (Alonso & Matos, 2008)

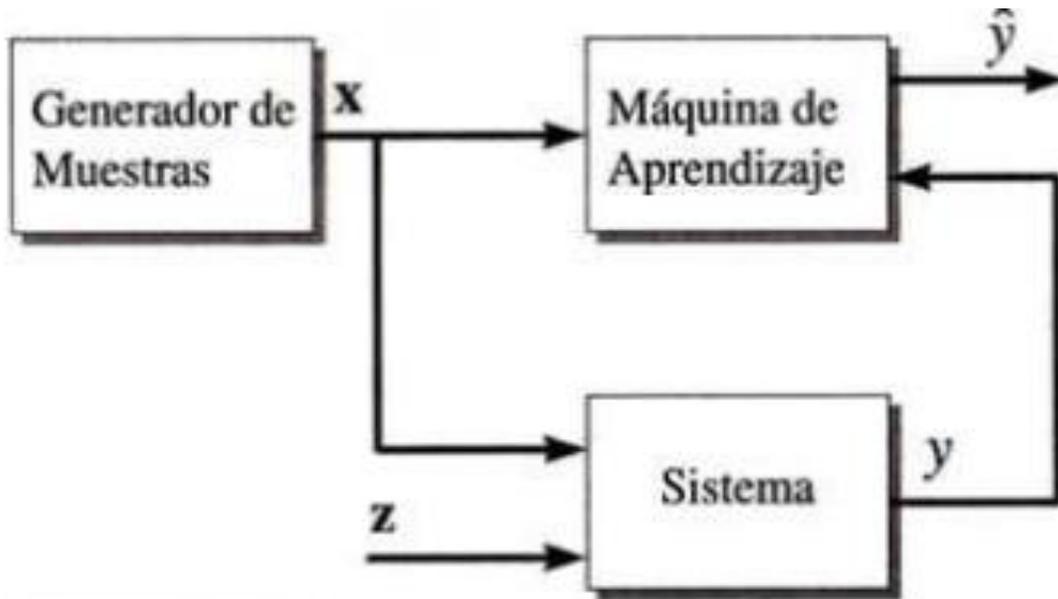


Figura 4. Una máquina de aprendizaje.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

2.5 PROCESAMIENTO MORFOLÓGICO

Las herramientas y los procedimientos matemáticos que se utilizan para extraer los componentes de una imagen utilizados para la representación y descripción de regiones se denominan morfológicos de procesamiento o morfología. La Morfología ofrece soluciones a muchos problemas de procesamiento de imágenes, ya que propone operadores generales que manejan grupos de píxeles, que se interpretan como información en la imagen, que pueden ser texturas, formas, líneas de contorno, etc. La morfología puede aplicarse a imágenes de intensidad e imágenes binarias. Los principales operadores morfológicos para las imágenes binarias son erosionar y dilatar. Todos estos operadores se pueden describir simplemente en términos de aumentar o eliminar píxeles de la imagen binaria de acuerdo con ciertas reglas que dependen del patrón de vecindad de la imagen binaria. pixel de modificación. Para las imágenes binarias, los píxeles blancos son regiones de importancia, mientras que los píxeles oscuros son el fondo de la imagen. (Rong, Chen, Liu, & Haeusler, 2009)

2.5.1 DILATACIÓN

La dilatación es el operador morfológico complementario a la erosión. La dilatación agrega píxeles siguiendo un patrón similar a la erosión, cambia el valor de un píxel del fondo de la imagen a 1 binario si toca un píxel que es parte de una región. Esto agrega una capa de píxeles en la periferia de todas las regiones, lo que provoca un aumento en las dimensiones y puede provocar que las regiones se unan. Los efectos gráficos del operador de dilatación

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

se pueden ver en la Figura 5. La figura original a la que se aplicó el operador de expansión es la figura 5 (a). (Alonso & Matos, 2008)

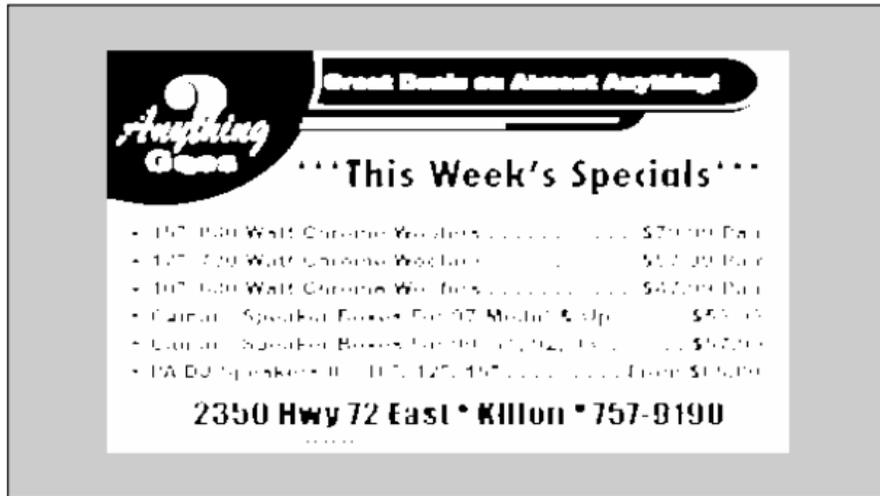


Figura 5. Imagen dilatada.

2.5.2 EROSIÓN

Erosión cambia los valores de los píxeles que originalmente eran 1 binario a 0 binario. Gráficamente, esto corresponde a la eliminación de los píxeles de la imagen. El propósito es eliminar los píxeles que no corresponden a ciertas regiones. Por ejemplo, cuando se usan escalas de grises en imágenes de intensidad, los píxeles se seleccionan de acuerdo con umbrales que establecen rangos de brillo, pero algunos píxeles no caen dentro de las regiones que tienen ese valor de brillo. Su valor de luminosidad puede ser debido al ruido finito de la imagen o porque se encuentran en un borde entre una región de luz y oscuridad y promediar el valor de brillo de la región seleccionada por el umbral. El operador de erosión elimina una capa de píxeles de la periferia de todas las regiones, lo que provoca una disminución de las dimensiones y puede crear otros problemas si esto hace que un contorno deje de ser continuo y romper en partes. La figura 6 muestra el cambio en una imagen debido a la erosión. (Alonso & Matos, 2008)

	<p style="text-align: center;">INFORME FINAL DE TRABAJO DE GRADO</p>	Código	FDE 089
		Versión	03
		Fecha	2015-01-22



Figura 6. (a) Imagen original. (b) Imagen erosionada

2.5.3 APERTURA Y CIERRE

La combinación de una erosión seguida de una dilatación se llama abrir o apertura, refiriéndose a la habilidad de esta combinación para abrir espacios entre elementos que apenas se tocan. Es una de las secuencias más utilizadas para remover el ruido de píxeles de imágenes binarias. La apertura empareja el contorno de una imagen, rompe prolongaciones estrechas y elimina salientes finas. (Alonso & Matos, 2008)

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

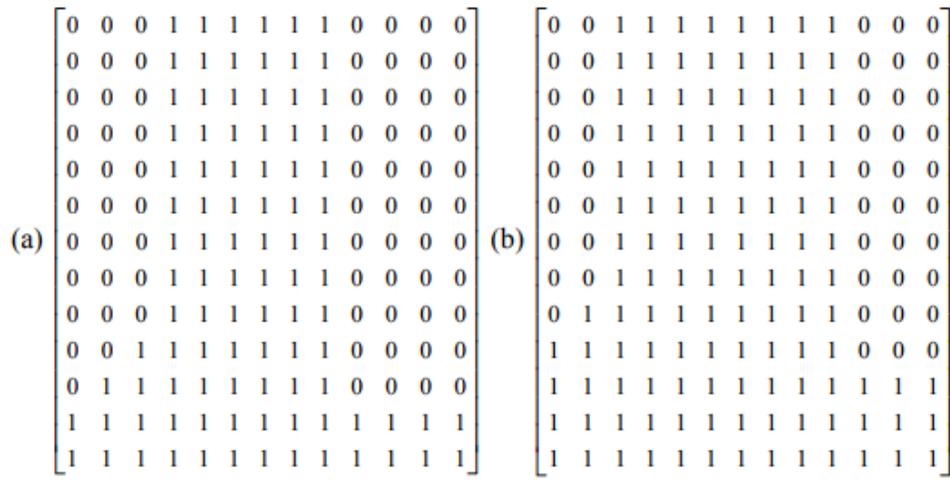


Figura 7. (a) Arreglo extraído de la imagen original. (b) Arreglo extraído de la matriz obtenida tras el proceso de dilatación.

Hay varios parámetros que se pueden ajustar para modificar la erosión y la expansión, particularmente el patrón de vecindad y el número de iteraciones. En la mayoría de las operaciones de apertura, éstas se conservan las mismas para la erosión y la expansión. Los efectos del operador de apertura en la imagen de referencia se observan en la Figura 8. (Alonso & Matos, 2008)



Figura 8. Imagen procesada con el operador de apertura.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

Cambiar el orden de la dilatación y la erosión cambia el resultado, si la dilatación se lleva a cabo en primer lugar, en lugar de eliminar los píxeles aislados que están en 1, el resultado es llenar los lugares donde los píxeles aislados están en 0, perdiendo los píxeles dentro de los objetos o en bandas estrechas entre porciones de objetos.

Cerrar tiende a coincidir con secciones de los contornos, pero, en oposición a la apertura, por lo general fusibles estrechos y estrechas entradas delgadas, elimina pequeños agujeros y llena huecos en el contorno. Los parámetros de expansión y erosión también se equilibran como en la apertura. Los efectos sobre las imágenes del operador cercano se muestran en la Figura 9. (Alonso & Matos, 2008)



Figura 9. Imagen procesada con el operador “cerrar”.

2.5.4 FILTROS DETECTORES DE BORDES

Los filtros de realce de borde asignan valores de nivel de gris a los píxeles de la escena original, dependiendo de la influencia de sus píxeles vecinos. La mayor o menor influencia de la proximidad del píxel central; será una función de valores positivos, nulos o negativos determinados por el usuario y atribuidos a los elementos de la matriz estructuradora, según la configuración del filtro utilizado. Por lo tanto, a través de la combinación de estos valores de entrada, obtendrá un mayor o menor mejora de la escena. (Coca, 2008)

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

Las máscaras de filtraje que se presentan a continuación:

$$\begin{array}{cc}
\begin{vmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{vmatrix} & \begin{vmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{vmatrix} \\
\text{Grisés Altos} & \text{Grisés Medios}
\end{array}$$

Figura 10. Mascaras de filtros.

2.6 ECUALIZACIÓN DE HISTOGRAMAS

La ecualización de histogramas es una técnica frecuentemente usada en el procesado digital de imágenes, para mejorar el brillo y el contraste, y optimizar el rango dinámico de la escala de grises. De manera simple corrige automáticamente las imágenes demasiado brillantes, demasiado oscuras o con poco contraste. Los valores de los niveles de gris se ajustan dentro de un determinado rango y la entropía de la imagen se maximiza.

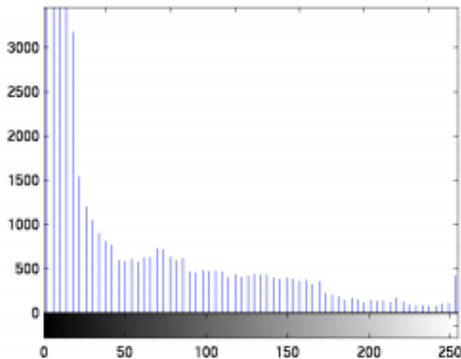
- $H'(i) = \sum_{0 \leq j < i} H(j)$ (5)
- Imagen Ecualizada(x,y) = H'(imagen(x,y)) (6)

Donde h es el histograma y h' es distribución acumulada.

Es una forma de manipulación de histogramas que reduce automáticamente el contraste en las áreas muy claras o muy oscuras de una imagen. También expande los niveles de gris a lo largo de todo intervalo. Consiste en una transformación no lineal que considera la distribución acumulativa de la imagen original, para genera una imagen resultante cuyo histograma será aproximadamente uniforme.

Sin embargo, se consigue una aproximación al dispersar los picos del histograma de la imagen, dejando intactas las partes más bajas. Este proceso se obtiene a través de una función de transferencia que tiene una alta inclinación siempre que el histograma original presenta un pico y una baja inclinación en el resto del histograma. (Sariyanidi, 2012)

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22



Figura(a). Contraste oscuro

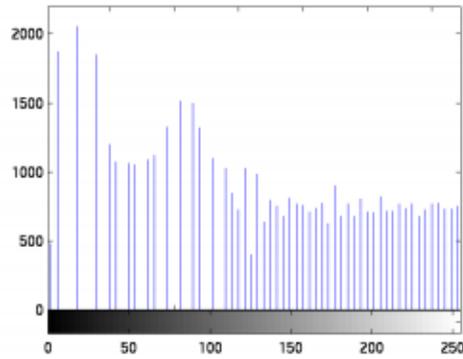


Figura (b). Contraste distribuido

Figura 11. Ecuilización de Histogramas.

2.7 MÉTODO RECONOCIMIENTO FACIAL EN CASCADA

2.7.1 DETECCIÓN DE ROSTRO

Debe quedar claro que la detección de una cara es más simple y requiere menos trabajo que determinar la cara a la que pertenece la persona. Para poder determinar que una imagen contiene una cara (o varias) necesitamos ser capaces de definir la estructura global de una cara. Aunque una cara difiere mucho de las otras, por lo general las caras humanas contienen similitudes entre sí que pueden facilitar la detección de una cara como son: nariz, ojos, frente, barbilla y boca; y todos estos constituyen la estructura general de una cara. (Guerrero, 2014)

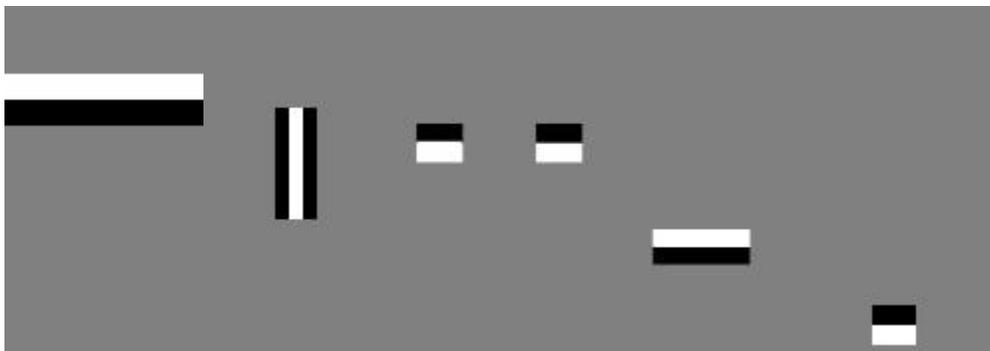


Figura 12. Características de un Rostro.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

Cada una de estas figuras representa una característica general de un rostro humano. La combinación de todas las características juntas se asemeja a una cara.



Figura 13. Características de un Rostro unificadas.

Al determinar si cada una de estas características es similar a alguna parte de la imagen, se puede concluir si dicha imagen contiene una cara o no. Hay que tener en cuenta que esto no tiene por qué ser una coincidencia exacta; sólo se tiene que saber si cada una de estas características corresponden a una parte de la imagen. La técnica utilizada para este propósito es el emparejamiento de plantillas. (Coca, 2008)



Figura14. Ejemplo detección de Rostro.

Actualmente existen algoritmos para detectar rostros que dependen de los escenarios a considerar. Una posible clasificación de los algoritmos de detección facial es: Métodos

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

basados en la Facial: buscan encontrar las características presentes en cualquier cara: ojos, cejas, labios, boca, barbilla, líneas de contorno, etc. sintetizar un modelo a partir de un conjunto de imágenes de entrenamiento. Trabajan con toda la imagen o una región de la misma sin buscar rasgos faciales localizados. (Coca, 2008)

2.7.4 RECONOCIMIENTO FACIAL

Es un método que ha ganado cierta popularidad dentro del sector de la seguridad para el control de acceso y para la identificación de personas, el gran problema con el que han tenido que lidiar con los diferentes algoritmos desarrollados, es que a diferencia del iris o la huella digital la cara puede ser alterado, además debe siempre cuidar la iluminación, el ángulo de donde se toma, las expresiones faciales, el uso de lentes y el tipo de cámara utilizada. En este tipo de sistemas se pueden encontrar dos errores principales denominados: FAR y FRR (Abraham, 2009, pág. 44).

La primera tiene la probabilidad de que el sistema autorice a una persona que no esté en el sistema y la segunda de impedir el acceso a una persona registrada en el sistema. Hay dos métodos básicos para el reconocimiento facial: el primero se basa en imágenes 2D, estos tipos de sistemas de tomar una fotografía de la cara y medir la distancia entre las diferentes características de la cara para crear una plantilla de identificación, el segundo método recientemente creado es la ventaja de este tipo de modelo en comparación con 2D es que elimina la posibilidad de que la imagen que se toma sea una fotografía en papel, la desventaja de este tipo de técnica es que el procesamiento para crear el modelo 3D requiere más hora. (Jain, Flynn, Abraham, 2009, págs. 44-45):

Reconocimiento de Caracteres: Este método se basa en relaciones geométricas tales como áreas, distancias y ángulos que pueden existir entre las características particulares de la cara.

Reconocimiento de Apariencia: Esta técnica utiliza las propiedades globales de la imagen de la cara para determinar los patrones de verificación, esto se hace usando vectores procesados por una computadora para representar la cara eficientemente. Este método es

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

el más utilizado dentro de la industria biométrica. A continuación, se presentan los algoritmos más populares (Abraham, 2009, págs. 45-70)

el Análisis de Componentes Principales (PCA).

Análisis de Discriminación Lineal (LDA).

Análisis de Componentes Independientes (ICA).

Análisis de características locales (LFA).

Reconocimiento de múltiples coincidencias (EBGM).

redes neuronales (NN).

Máquinas de Soporte Vectorial (SVM).

En conclusión, las técnicas para realizar el reconocimiento de la cara han evolucionado significativamente, considerando todos los problemas que han tenido que atender al crear un algoritmo, este tipo de biometría puede ser utilizado en conjunción con otros tipos de técnicas, como el reconocimiento del iris, para hacer la seguridad más eficiente.

2.7.3 EIGENFACES Y FISHERFACES

Eigenfaces: Al conjunto de los auto vectores resultantes se les llama eigenfaces, cuya traducción literal al español sería "propias caras"; sin embargo, como no hay traducción oficial, a través de este documento se les llamará eigenfaces. Este término tiene su origen en que, si cambian el tamaño de los eigenfaces al tamaño de las imágenes de entrenamiento y se expanden, por ejemplo, su valor mínimo es cero y el máximo es 255 (valores mínimo y máximo de un píxel para una imagen 8 -bit), cuando se ven como imágenes actuales, se pueden ver caras. La figura 15 muestra un conjunto de entrenamiento formado por las primeras 10 imágenes de cada individuo de la base de datos de la Universidad de Yale (15 individuos) [21]. La Figura 16 y la Figura 17 muestran respectivamente la imagen media y los 20 primeros eigenfaces de ese conjunto calculados en MATLAB. La Figura 17 también muestra el porcentaje de la varianza acumulada total que explica cada uno de los primeros 20 eigenfaces. Por lo tanto, para conocer el porcentaje de varianza explicado por uno de los

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

eigenfotos en particular, el porcentaje de varianza acumulada de este eigenface debe ser restado del anterior. (Ballesta, Gil, & Reinoso, 2010)



Figura15. Set de entrenamiento formado por algunas de las imágenes de la base de datos de Yale University [21].



Figura 16. Imagen media del set de entrenamiento de la Figura 15. Esta figura fue obtenida utilizando MATLAB.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22



Figura 17. Veinte primeras eigenfaces del set de entrenamiento de la Figura 15. Esta figura fue obtenida utilizando MATLAB.

2.7.4 HISTOGRAMAS DE PATRONES BINARIOS LOCALES

El método LBPH proporciona un operador de análisis de textura que se define como una medida de la textura en una escala de gris invariable, derivada de una definición general de la textura por los vecinos locales. La forma actual del operador LBP es muy diferente de su versión básica: la definición original se extiende a un conjunto de vecinos arbitrarios circulares y se han desarrollado nuevas versiones, sin embargo, la idea principal es la misma: un modelo binario de código que describe el patrón de textura local que se construye por el umbral de un conjunto de vecinos por el valor gris de su centro. (Monadjemi, 2011) El operador tiene que ver con muchos otros métodos conocidos de análisis de textura en la Figura 18. Observe las relaciones del operador LBP con los otros métodos.

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

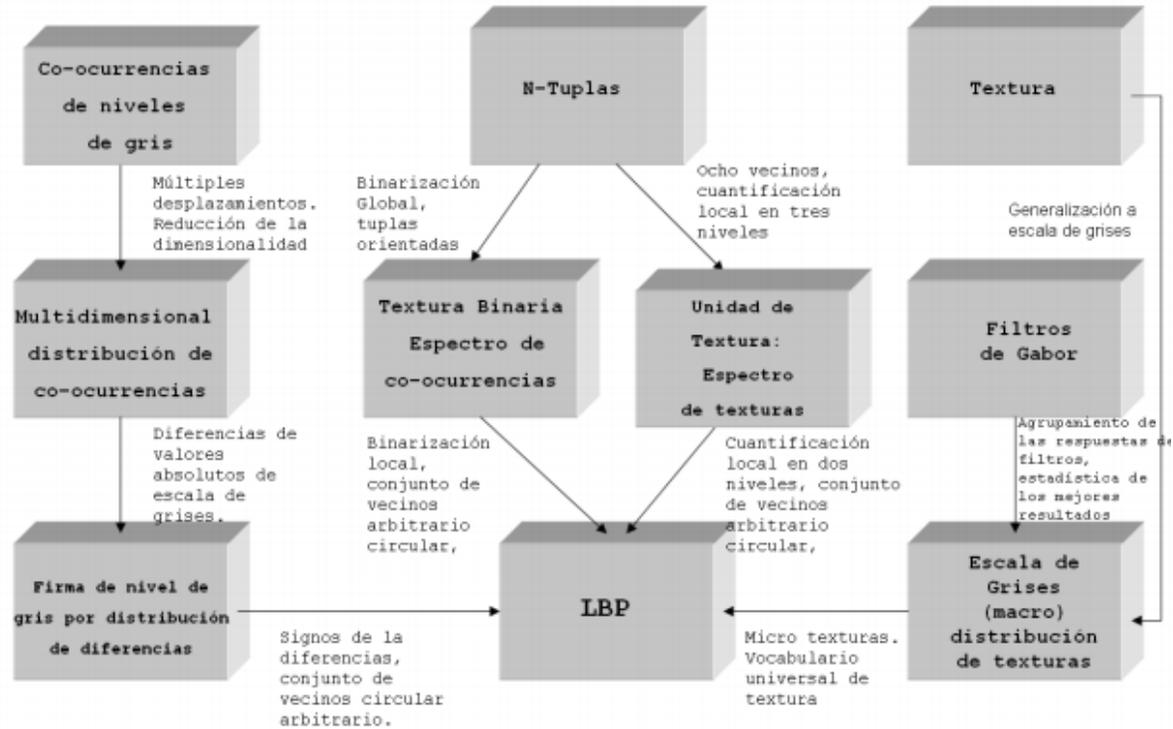


Figura18: LBP operador en el campo de análisis de texturas

2.8 DETECCIÓN FACIAL CON PUNTOS

para la ubicación de los puntos característicos en la cara usamos Active Shape Model (ASM) que crea un modelo de las características que uno desea encontrar. Este modelo se define por una serie de puntos de modelo, así como la conexión entre los puntos. Examine la región alrededor de cada punto encontrado en la imagen para encontrar la mejor "opción" en la imagen que se ajusten a los otros puntos. Mueva el modelo de punto a punto para que ubique los otros puntos de la imagen de una manera apropiada. La adaptación se realiza utilizando los bordes de la imagen. En la figura 19 vemos cómo se extraen algunas características de la imagen a tratar, posteriormente estas características serán analizadas por el modelo ASM para desentrelazar si ese objeto es una cara. (Jian-zheng, 2011)

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22



Figura 19. Extracción de características de una imagen tras haber detectado previamente el área del rostro.

2.8.1 SURF

Es un detector y un descriptor de las características locales que se pueden utilizar para las tareas tales como reconocimiento del objeto, registro, clasificación o reconstrucción 3D. El detector localiza los puntos de interés en una imagen y el descriptor describe las características de los puntos de interés; a partir de la cual se basa para construir vectores de características. Las características son invariantes al desplazamiento, rotación, escalamiento y también son parcialmente invariantes a la iluminación. (Ballesta, Gil, & Reinoso, 2010)

El descriptor de SURF representa una distribución de las respuestas de Haar-wavelet en el entorno de punto de interés y utiliza imágenes integrales eficientemente. En este trabajo se han estudiado las tres versiones de este descriptor: El descriptor estándar de SURF, que tiene una dimensión de 64, la versión extendida (E-SURF) con 128 elementos y la versión vertical (U-SURF). La versión U-SURF no es invariante a la rotación y el vector tiene 64 elementos (Bay et al., 2006).

2.8.2 HOMOGRAFÍAS

Las homografías son transformaciones proyectivas que establecen una correspondencia entre elementos de la misma especie. Esto significa que en un punto corresponde otro punto y una recta corresponde a otra recta. Es la correspondencia punto a punto y recta a recta de dos figuras planas. Por lo tanto, las transformaciones como la traslación o la

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

homotecia son homografías. El tipo más simple de formas homográficas son dos secciones planas de la misma radiación.

Cuando a un objeto se le aplica una fuente luminosa éste genera una sombra, entre el objeto y su sombra existe una relación biunívoca, de manera que a cada punto del objeto le corresponde otro en su sombra y viceversa. Así pues, se ha establecido una relación de transformación.

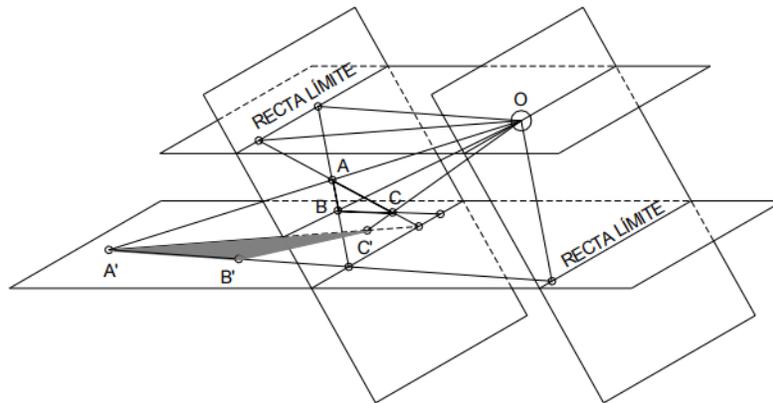


Figura 20. Transformación proyectiva.

Son transformaciones proyectivas las aplicaciones de un espacio lineal (rectas, planos, espacios tridimensionales) en otro, de manera que cuatro puntos en línea recta se transforman en cuatro puntos en línea recta. Las transformaciones pueden ser de tres tipos: ISOMÉTRICAS: También llamadas movimientos. Son aquellas que conservan las medidas de los segmentos y de los ángulos de la figura original y su transformada (TRASLACIÓN, GIRO y SIMETRÍA).

ISOMÓRFICAS: Son aquellas que conservan las formas. Se pueden establecer relaciones de proporcionalidad entre dos figuras transformadas (HOMOTECIA Y SEMEJANZA).

ANAMÓRFICAS: Son las transformaciones que no conservan las formas (HOMOLOGÍA y AFINIDAD).

Lo que es normal en el estudio homográfico de los planos mencionados, entre los cuales se establece la proyectividad, debe hacerse con ellos superpuestos, de modo que se presenta

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

la posibilidad de elementos dobles. El número y distribución de los elementos dobles establece los diferentes homógrafos. Una línea doble no está compuesta de puntos dobles, pero, a lo sumo, tendrá dos, si hay más, será una identidad, una estructura compuesta enteramente de puntos dobles, y recibe comúnmente el nombre de eje. Tampoco se duplica todas las líneas que pasan a través de puntos dobles, pero, a lo sumo, habrá dos, en caso de que haya más de dos líneas dobles que pasen por el punto, se enfrentará con el caso en que el punto es centro. Es también una declaración verificable y demostrable que establece que en toda homografía entre planos diseminados (o regulados), si hay un eje, hay un centro y viceversa.

Para poder calcular la matriz que transforma la imagen es necesario obtener al menos cuatro correspondencias entre los puntos x y x' no alineados de ambas imágenes, la que queremos transformar y a la que queremos que se transforme. Sin embargo, en la práctica, para conseguir mayor robustez en el cálculo de la homografía se obtiene un número elevado (decenas o incluso cientos) de correspondencias y se busca la homografía que mejor se adapta a todas ellas. La técnica para obtener los puntos característicos junto con sus descriptores ha sido SURF. Una vez aplicado SURF a ambas imágenes se hace un matching entre los puntos característicos de ambas imágenes, el cual es el resultado de aplicar un umbral a la distancia euclidiana entre sus descriptores (Rong, Chen, Liu, & Haeusler, 2009).

2.8.3 DETECTORES DE PUNTOS DE INTERÉS

La detección de puntos de interés corresponde a la primera etapa durante el procesamiento de una imagen, que implica la individualización de puntos, que por sus características ayudan en su conjunto a definir el objeto contenido en una imagen. Se pretende que los detectores sean eficientes con respecto a los recursos computacionales que utilicen, esto influye directamente en la cantidad de fotogramas por segundo que se pueden procesar y en las plataformas de hardware sobre las que se puede implementar el procesamiento en tiempo real en caso de que de ser necesario. (Guerrero, 2014)

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

2.8.4 DESCRIPCIÓN DE PUNTOS DE INTERÉS

Un detector ideal de puntos de interés localiza dichos puntos repetidamente a pesar del cambio de punto de vista siendo, fiable ante transformaciones de la imagen. Un descriptor ideal obtiene de cada punto de interés información relevante y distintiva de la región que lo rodea, de manera que la misma estructura pueda ser reconocida si es encontrada en otra imagen. Implica un verdadero desafío describir de manera eficiente puntos de interés en forma estable y compacta, con representaciones robustas invariantes a escala, rotación, diversas transformaciones y el ruido. (Guerrero, 2014)

2.9 COMPUTACIÓN COGNITIVA

La computación cognitiva es la simulación de procesos de pensamiento humano en un modelo computarizado. La computación cognitiva, o cómputo cognitivo, implica sistemas de autoaprendizaje que utilizan la minería de datos, el reconocimiento de patrones y el procesamiento del lenguaje natural para imitar el funcionamiento del cerebro humano. El objetivo de la computación cognitiva es crear sistemas informáticos automatizados que sean capaces de resolver problemas sin necesidad de asistencia humana.

La computación cognitiva es la simulación de procesos de pensamiento humano en un modelo computarizado. La computación cognitiva, o cómputo cognitivo, implica sistemas de autoaprendizaje que utilizan la minería de datos, el reconocimiento de patrones y el procesamiento del lenguaje natural para imitar el funcionamiento del cerebro humano. El objetivo de la computación cognitiva es crear sistemas informáticos automatizados que sean capaces de resolver problemas sin necesidad de asistencia humana.

Los sistemas computacionales cognitivos usan algoritmos de aprendizaje automático. Dichos sistemas adquieren continuamente conocimientos a partir de los datos alimentados en ellos mediante la minería de datos para obtener información. Los sistemas refinan la forma en que buscan patrones, así como la forma en que procesan los datos para que sean capaces de anticipar nuevos problemas y modelar posibles soluciones.

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

La computación cognitiva se utiliza en numerosas aplicaciones de inteligencia artificial (IA), incluyendo sistemas expertos, programación en lenguaje natural, redes neuronales, robótica y realidad virtual. El término computación cognitiva está estrechamente asociado con el sistema cognitivo de IBM, Watson. (Kelly, 2015)

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

3. METODOLOGÍA

3.1 ESQUEMA GENERAL DE LAS ETAPAS Y BASE DE DATOS

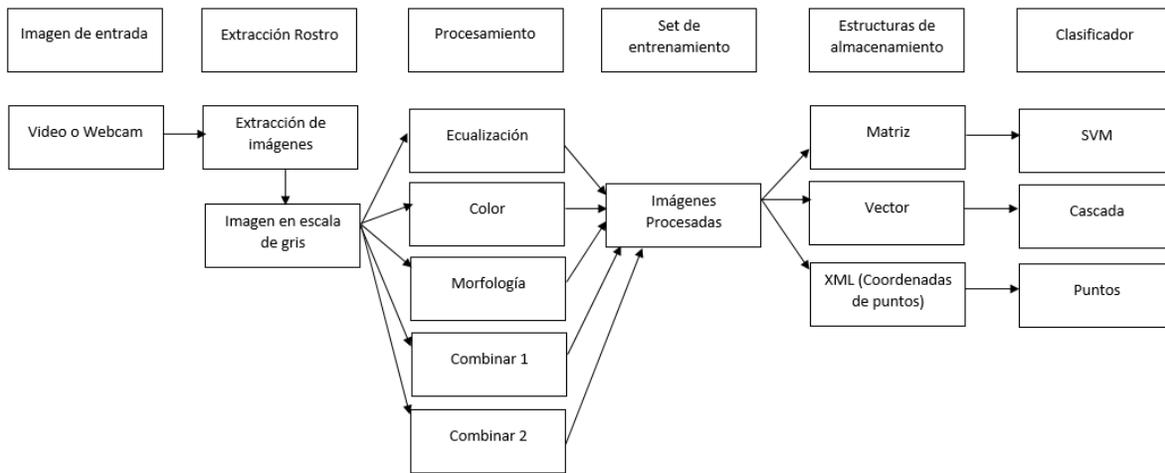


Figura 21. Esquema General.

Debe mencionarse que cada proceso realiza combinaciones diferentes (ver sección 3.3) que dan lugar a imágenes procesadas, que corresponden a dichos procesos realizados, es decir, si se ha realizado un procesamiento que implica morfología; las imágenes del conjunto de entrenamiento corresponderían al proceso morfología; esto implica que las imágenes que se generan no se mezclan con las imágenes de otros procesos, después de todo en este trabajo se aplicaron diferentes versiones que dan como resultado diferentes conjuntos de entrenamientos correspondientes cada uno a sus respectivos procesos. Finalmente se debe mencionar que las imágenes procesadas(rostros) son pasadas a una estructura dependiendo del clasificador y entonces el proceso de predicción o evaluación se lleva a cabo como se explica más adelante en cada uno de los respectivos métodos.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

3.1.1 EXTRACCIÓN DE IMÁGENES:

La extracción de imágenes consiste en el uso de vídeos con la misma cantidad de tiempo, aproximadamente de 10 segundos con una buena atmósfera iluminada y con fondo blanco para reducir la cantidad de información; también consiste en el uso de archivos XML de apoyo que ayudan a encontrar el área de interés de la región de la cara; una vez que se reconoce la región de interés, se recorta la imagen de modo que todos tengan el mismo tamaño (60x60) o 499x499 (detección de puntos), con el fin de reducir el tiempo de procesamiento cuando se utilizan diferentes metodologías. Debe mencionarse que el resultado final es una imagen del rostro en escala de grises.

3.2 METODOLOGIAS

Cabe mencionar que, durante el desarrollo de las metodologías, cada método utilizado tanto para SVM como para clasificador en cascada son multiclase.

3.2.1 MÁQUINAS DE VECTORES DE SOPORTE O SVM:

El desarrollo de SVM consistió principalmente en la extracción de imágenes de cara y la ecuilización del histograma de cada una de ellas, con el fin de distribuir mejor las escalas de grises y obtener un mayor contraste. Luego se decidió probar diferentes tratamientos en el procesamiento de imágenes para luego combinarlos con el modelo SVM; esto para poder concluir cuál es el mejor tratamiento de imágenes para las diferentes metodologías basadas en los resultados obtenidos, entre ellos se realizó una aplicación de color; que es la transformación del color BGR a otros colores como XYZ o HLS; que fueron las que permitieron una mejor separación de la cara en comparación con las otras conversiones de color realizadas en el presente proyecto. Estos se combinan con una aplicación de mapa de color Hot que se convierte a escala de gris, lo que permite mejores resultados con respecto a los tonos de piel. Otros tratamientos que se llevaron a cabo fueron la aplicación de filtro para detectar el contorno de la cara de la persona por medio de Sobel y su gradiente; Los

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

otros tratamientos fueron combinaciones de los dos tratamientos mencionados anteriormente.

Para la elección de modelos de color en las imágenes, se implementó la implementación y evaluación de varias transformaciones, esto se realizó para determinar cuál de estas transformaciones fue la más adecuada para utilizar con las caras de diferentes tonos de color de la piel. Como se muestra en las figuras 22 y 23, se encierra en un recuadro las que facilito la detección del rostro y que fueron utilizadas en el presente proyecto (XYZ y HLS). Por último, cabe mencionar que el modelo elegido en la mayoría de las pruebas y con el que se realizó el análisis de resultados fue el modelo XYZ.

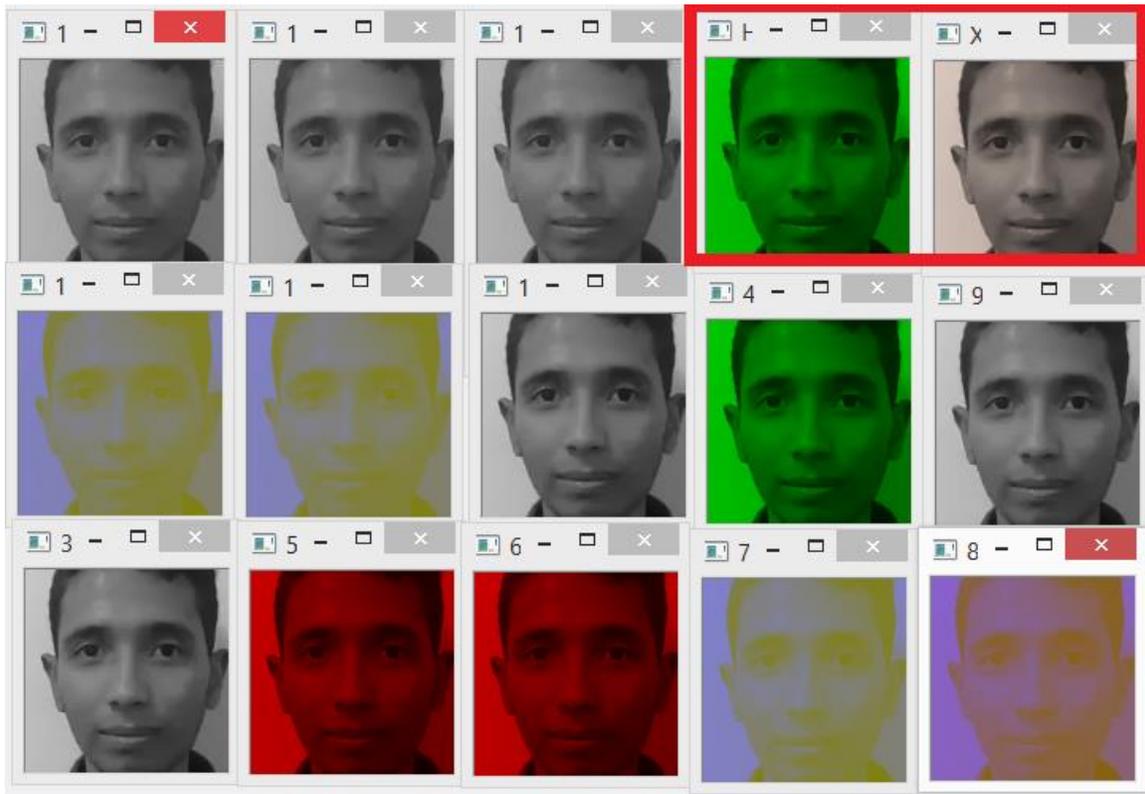


Figura 22. Transformada de varios colores.

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

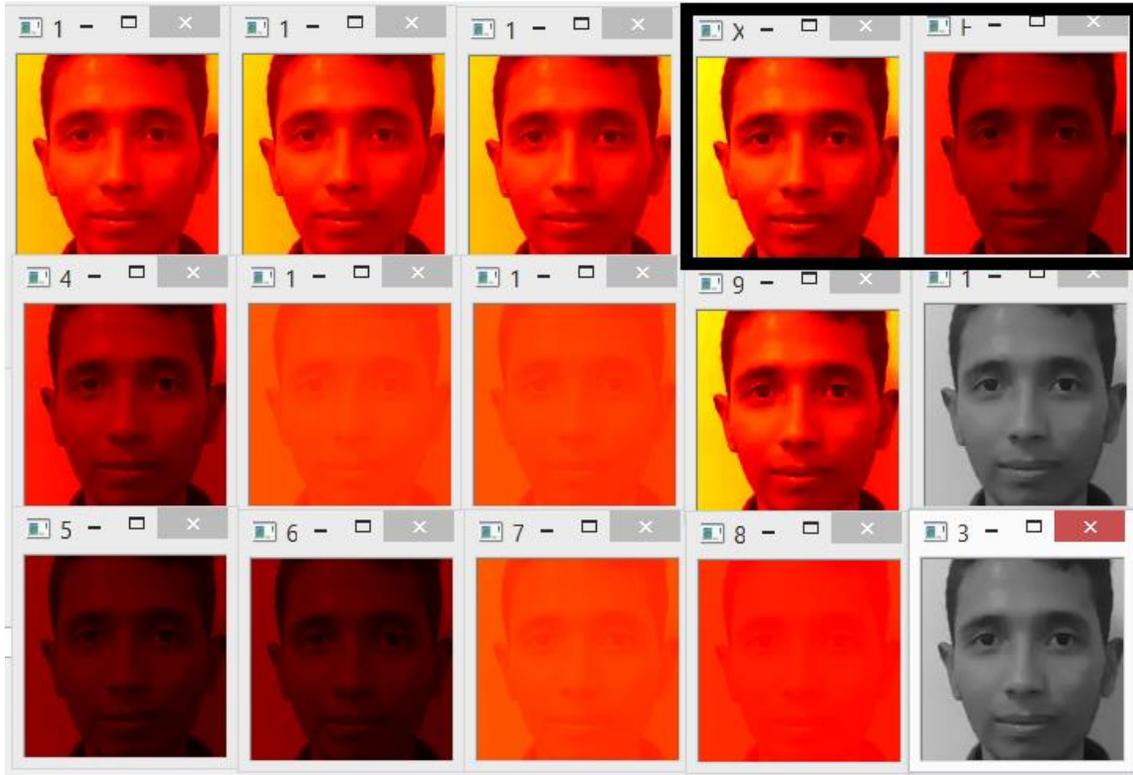


Figura 23. Aplicando mapa de color Hot en las transformadas.

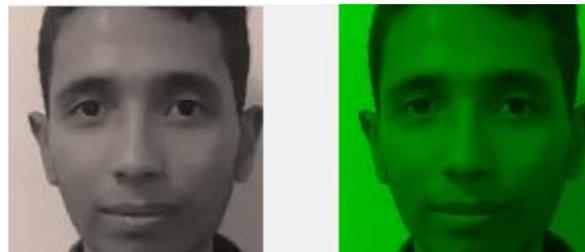


Figura 24. Color XYZ y HLS.



Figura 25. Mapa de color1 (Hot) Aplicado.

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22



Figura 26. Mapa de color2 (Rainbow) Aplicado.



Figura 27. Gris Color1 (Ecuilizada).



Figura 28. Gris Color2 (Ecuilizada).



Figura 29. Gris color1 (Ecuilizada y Sobel).

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22



Figura 30. Gris color2 (Ecuilizada y Sobel).



Figura 31. Gris color1 (Ecuilizada y Sobel, versión float).



Figura 32. Gris color2 (Ecuilizada y Sobel, versión float).

Una vez realizadas todas las transformaciones, se aplicó el modelo SVM proporcionado por Opencv, el cual fue entrenado con alrededor de 200 imágenes de ocho personas, las cuales se encuentran en posición frontal y realizan diferentes gestos; la capacitación del modelo se llevó a cabo utilizando diferentes funciones con el fin de lograr una mejor separación y obtener un mejor resultado estadístico al predecir y realizar el reconocimiento facial. Este entrenamiento consistió en llevar todas las imágenes a una matriz, de manera que cada una de las imágenes se almacena en las filas de dicha matriz, además se crea un vector de columnas de etiquetas que indica en cada fila a cada persona a la que pertenecen las

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

imágenes; estas matrices se envían al modelo SVM de tal manera que lo que se almacena en los vectores son las caras con sus respectivas transformaciones y junto a sus respectivas etiquetas. Una vez finalizado este almacenamiento, se generó un archivo XML correspondiente al conjunto de imágenes, de este modo se generaron los vectores que permiten las predicciones de las diferentes imágenes obtenidas en tiempo real, cabe destacar que la generación del archivo XML dependía de las funciones del modelo, de los parámetros de la misma y de los tratamientos ya mencionados.

El modelo SVM utiliza el archivo XML generado para realizar la predicción y reconocimiento de las caras, este proceso se realizó mediante una cámara diferente de la del ordenador, ya que quería obtener una imagen más detallada, a la que se realizaron los mismos pasos de procesamiento que en la etapa de entrenamiento; terminado estos procesos se evalúa la imagen obtenida de la cámara y procesada con el modelo SVM cargado del archivo XML, esta imagen se compara con los vectores y finalmente se obtiene un resultado que indica una etiqueta y ésta indica a la persona, posteriormente se muestra en tiempo real el resultado obtenido, vale la pena mencionar que el modelo utiliza una función lineal (basada en los vecinos más cercanos) de las muchas funciones probadas, esto porque la función lineal obtuvo mejores resultados, además de que el tiempo de generación de los archivos XML es significativamente menor en comparación con las otras funciones. Este proceso se repite cada frame de vídeo de tal manera que se genera un reconocimiento en cada uno de ellos; esto indica que se obtendrá un porcentaje de reconocimiento porque el modelo no genera un 100% de predicción y al mismo tiempo variará según ciertas condiciones, por lo cual, lo que se quiere es tener un buen porcentaje de acierto para poder hacer los comparativos con los demás métodos de reconocimiento facial.

El modelo SVM que proporciona opencv no permite un re-entrenamiento incremental basado en el archivo XML generado en el entrenamiento, por lo cual a la hora de ingresar una nueva persona se necesita cambiar el modelo, lo que también implica cambiar el dataset, por lo que toma más tiempo ya que todos los pasos deben ser generados de nuevo, desde la extracción de nuevas imágenes, el procesamiento de dichas imágenes y finalmente

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

la generación del nuevo archivo XML, además también vale la pena mencionar que una vez generado el nuevo archivo XML, el porcentaje de aciertos del reconocimiento facial puede variar con respecto al archivo anteriormente generado.

3.2.2 MÉTODO CASCADA:

El desarrollo de este segundo método tuvo en cuenta las transformaciones, procesos y extracciones de imágenes del método anterior, además del uso de tres modelos diferentes, como el modelo de Fisher, el modelo de Eigen y el modelo LBPH; esto con el fin de obtener el mejor resultado para poder hacer la comparación con los otros métodos. Como con el primer método, se realizan los mismos procesos y transformaciones de imagen. Una vez que este paso se ha terminado, las imágenes se llevan a un elemento estructural que almacena esas imágenes, que a diferencia del primer método es sólo un vector de filas que las contiene y no una matriz; esto se hace para simplificar el uso de clasificadores en cascada al enviar los parámetros necesarios para generar el archivo de predicción. También vale la pena mencionar que también se generaron etiquetas que identifican qué clase o nombre pertenece a la imagen de la cara de la persona, también almacenada en un vector de filas. En base a lo anterior y ya teniendo las estructuras que han almacenado las imágenes y etiquetas, el entrenamiento de los tres modelos se genera. Para ello se utilizaron las imágenes, etiquetas, una cantidad de componentes de 80, más un umbral de confianza también de valor 80; esta cantidad de componentes y valor umbral permitió la mejora en los resultados, además de ser un poco más estricto cuando se trata del reconocimiento de la cara de la persona. La finalización del entrenamiento de los modelos da como resultado seis archivos, cada uno correspondiente a los tres modelos, pero con dos versiones, es decir, una versión que contiene el umbral con el valor ya mencionado y la otra versión sin umbral, esto se hizo para poder hacer una comparación entre el uso o no de los umbrales y de esta manera poder determinar si hay un cambio significativo en el tiempo de uso o no dicho umbral; los archivos generados contienen la información generada durante el entrenamiento. Que se utilizan para la evaluación y predicción de las imágenes; estas

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

imágenes siguen los mismos pasos de extracción y procesamiento, las cuales posteriormente se comparan en cada frame con el archivo generado por el modelo. Finalmente, se repitió todo el proceso anterior para cada uno de los archivos generados durante el entrenamiento y así concluir cuál es el mejor modelo, además si es mejor con umbral o sin umbral. Debe recordarse que el re-entrenamiento de este método se comporta de forma similar al método SVM, es decir, para poder hacer esto es necesario cambiar el conjunto de datos que resulta en un nuevo archivo que puede variar con respecto al anteriormente generado.

3.2.3 MÉTODO DETECCIÓN POR PUNTOS (Homografía):

En este método, a diferencia de los otros métodos, no necesita tanto procesamiento para poder llevar a cabo el entrenamiento, gracias a que este método es flexible con respecto a los ángulos de posición y con respecto a la iluminación del entorno; de modo que después de extraer la imagen y aplicar las transformaciones respectivas, se continuó con la extracción de puntos de los rostros de las personas mediante el uso del método SURF de opencv; para estas extracciones se hizo uso de un umbral con un valor de 300, el cual es recomendado por el mismo opencv, puesto que esto da más que suficientes puntos para trabajar. El método también necesita tanto las imágenes del modelo como las imágenes entrantes para poder identificar sus puntos clave o puntos junto a sus descriptores; ya teniendo los puntos de la imagen modelo como los de la imagen entrante, el método puede hacer una comparación por medio de la distancia euclidiana y por medio de los descriptores tanto de la imagen modelo como de los descriptores de la imagen entrante, generando así varios puntos para cada vecino utilizado en la comparación, que era una cantidad de dos vecinos, ya que con más vecinos la información se altera porque se generan muchos puntos que pueden alterar el resultado de la homografía. Con el fin de simplificar el número de puntos generados, se utilizó un filtro, mediante una máscara igual al tamaño de los descriptores entrantes que generaron una igualdad o coincidencia con los descriptores de la imagen del modelo; esta máscara redujo el número de puntos falsos dejando sólo

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

aqueellos que tenían alguna similitud o unicidad con la imagen del modelo, al igual que la máscara se aplicó de nuevo, pero esta vez buscando la orientación y el tamaño de los puntos, eliminando aquellos puntos que no cumplían con esta condición. Por último se reduce el número de puntos falsos o puntos que pueden alterar el resultado, el modelo puede generar la homografía que cumple las condiciones de las máscaras anteriores, es decir con la condición de coincidencia o unicidad junto con la condición de tamaño y orientación, dando como resultado la distancia de los puntos que son más similares entre la imagen entrante y la imagen del modelo; esto permite identificar qué imagen entrante corresponde a la imagen del modelo dando así una clasificación, finalmente lo que se hace es dibujar los puntos y estas distancias.

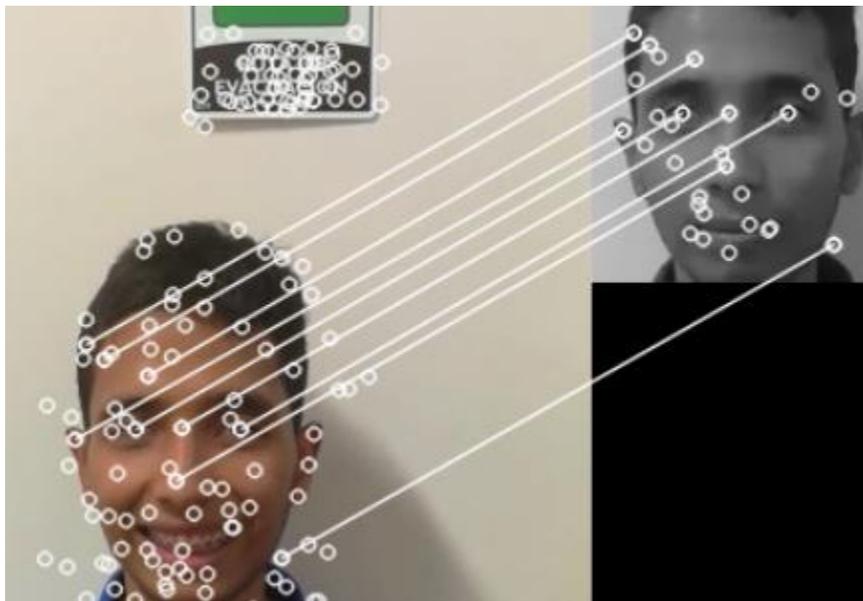


Figura 33. Ejemplo 1 Detección por puntos.

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

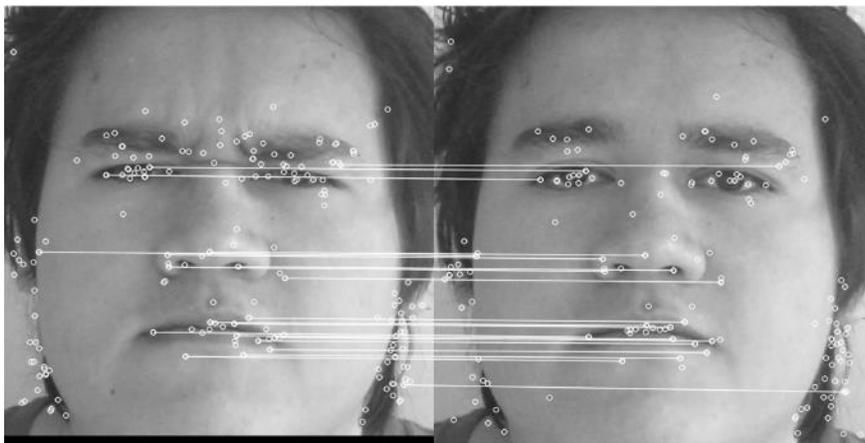


Figura 34. Ejemplo 2 Detección por puntos.

Cabe mencionar que las máscaras ayudan a prevenir los falsos positivos. El método en el momento de la predicción sólo necesita una base de datos de caras que iteran hasta el punto de encontrar la homografía; esto se hace por medio de la suma de puntos que están en cada imagen modelo, es decir que tienen varias imágenes modelo por persona, esa cantidad es arbitraria; la suma de puntos de cada una de las imágenes modelo correspondientes nos permite aproximar la clasificación correcta porque cuanto más similar es una imagen con la homografía más puntos identifica; en otras palabras, sumando todos los puntos de las imágenes modelo de la persona X se obtiene un valor alto en comparación con la suma de todos los puntos de las imágenes modelo de la persona Y, siempre que la imagen entrante tenga una homografía similar a la persona X, del mismo modo si la comparación entre la imagen entrante y la imagen modelo de la persona Y tiene una homografía similar, la suma de los puntos resultaría en un valor alto en Y que da como resultado la clasificación final de la persona, este método además permite que el re-entrenamiento sea fácil de realizar en comparación con los demás, ya que simplemente se tiene que añadir la cara de la nueva persona y no se necesita generar un nuevo archivo; una vez añadida la imagen de la cara de la nueva persona se repite todo el proceso explicado anteriormente.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

3.2.4 METODO DE DETECCION POR PUNTOS (sin homografía):

Esta metodología consistía primero en extraer las imágenes, pero a diferencia del método anterior, el resultado final no es una imagen en escala de grises. esta extracción de imágenes modelo consiste en primer lugar en identificar el área de interés de la cara, que pasa por varios procesos, como el cambio de color, la ecualización y el uso de la morfología, dando lugar a varias versiones de ellos; posteriormente a estos procesos se utiliza la función de soporte de opencv gfttdetector con un valor de radio de 1000 puntos y un tamaño de bloque de tres, estos puntos identificados gracias a esta función de soporte, se convierten en puntos clave o keypoints de la cara; estos puntos se almacenan en un vector de puntos y luego se dibujan en una estructura que contiene el mismo tamaño, canal y profundidad de la imagen del modelo. Por último, estos mismos puntos se pasan a una lista, por lo que se pueden guardar en un archivo XML que contiene su posición (x, y), así como su respectivo ángulo. Que, como se mencionó anteriormente, el resultado final no es una imagen en escala de grises, sino que es un archivo XML con los puntos que se extrajeron de cada imagen; vale la pena recordar que este proceso se realiza cada frame que puede variar, entre los 10 segundos que dura el video o hasta una cantidad deseada de imágenes.

Al evaluar las imágenes recibidas, es decir, las imágenes recibidas desde la webcam, se realiza un proceso igual al anterior, es decir, se extrae la imagen, se pasa por transformaciones que pueden ser color, morfología y ecualización, se extraen los puntos con la función de soporte de opencv. Una vez extraídos los puntos de la imagen entrante, se pasan a una lista y se cuentan con cada imagen del modelo, para comprobar las coincidencias o puntos de comparación entre cada imagen; por lo que cuanto más similar entre la imagen del modelo y la imagen entrante, más puntos se generarán. Esto permite realizar una vez que se termina la comparación con cada imagen del modelo; una suma de imágenes modelo con respecto a los puntos de cada una, que permite obtener el nombre de la persona del modelo que generó más puntos, este proceso es similar al recuento de puntos de la detección de puntos por homografía; donde cada imagen modelo tiene un número, es decir, se tiene una clase a, b, c, d, y así sucesivamente. con el mismo número de

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

muestras, se comparan con la imagen entrante lo que da como resultado una cantidad x de puntos, que dependen de la similitud; estas muestras se suman y la clase cuya suma de muestras genera una mayor cantidad de puntos es la ganadora, de esa clase se extrae la etiqueta que contiene el nombre de la persona. Este proceso se realiza cada frame, dando como resultado un nombre en cada uno de ellos. Del mismo modo, al método anterior de detección por puntos, el re-entrenamiento es fácil de realizar, ya que simplemente se necesita agregar el archivo XML de puntos de la nueva persona, el cual es un nuevo modelo que hace referencia a dicha persona agregada. Una vez hecho esto, se repite todo el proceso explicado anteriormente.

3.3 DIFERENTES TRATAMIENTOS DE LA IMAGEN

3.3.1 MÉTODOS DE COMBINACIONES

3.3.1.1 COMBINAR 1:

El método combinar 1 se encarga de procesar las imágenes mediante la extracción normalmente de ellas, luego se cambia de color bgr a xyz debido a que en las pruebas de transformación de color, esta arroja una mejor extracción de características, a la cual se le aplica un mapa de color hot, la razón es debido a que combinada con xyz, mejora las características; las imágenes se pasan a escala de gris y se ecualizan para mejorar el contraste, para posteriormente aplicar un sobel con su gradiente, con el fin de extraer los contornos de los rostros y eliminar información como el fondo; las imágenes resultantes tienen un valor float lo cual permite ver los contornos más claramente, debido a que no se pierde información cuando las imágenes se transforman a byte.

3.3.1.2 COMBINAR 2:

En combinar 2 se realizan los mismos pasos y procedimientos que se aplican en combinar 1, la excepción es que la imagen final cuyo valor es de tipo float se convierte a una imagen de tipo Byte para trabajar más fácil con ella a la hora de usar los métodos de reconocimiento y detección.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

3.3.1.3 MORFOLOGÍA:

Morfología consiste en extraer normalmente las imágenes, una vez extraídas dichas imágenes, se pasan a escala de gris y se ecualizan para mejorar su histograma y sus distribuciones de iluminación, una vez finalizado la ecualización se procede a sacar el gradiente de dicha imagen mediante una máscara tres x tres y mediante la función de apoyo de opencv, esto con el fin de agilizar los procesos y métodos.

3.3.1.4 COLOR:

La aplicación de color se basa simplemente en el proceso de extracción de las imágenes, de su transformación de color bgr a xyz, la combinación con el mapa de color Hot y finalmente su paso a escala de grises para posteriormente aplicarle una ecualización.

3.3.1.5 ECUALIZACIÓN:

El procesamiento que se realiza en este método como su nombre lo indica solo consiste en la extracción de imágenes, su paso a escala de grises y finalmente de su ecualización.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

3.3.2 DIAGRAMAS FLUJOS DE COMBINACIONES

3.3.2.1 COMBINAR 1:

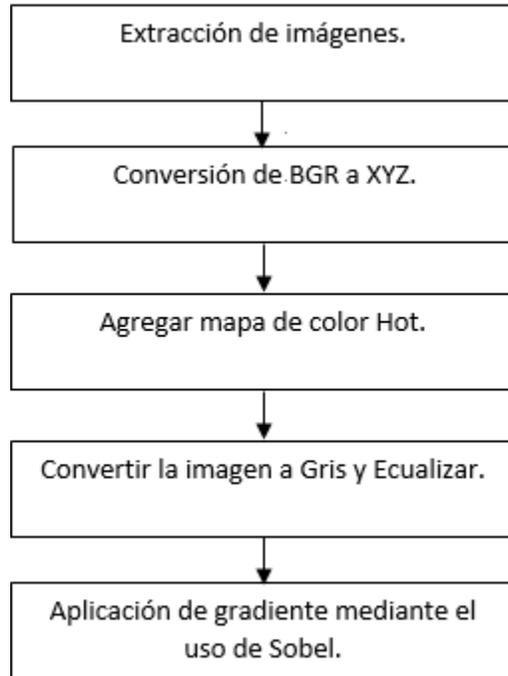


Figura 35. Flujo de proceso combinar1.

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

3.3.2.2 COMBINAR 2:

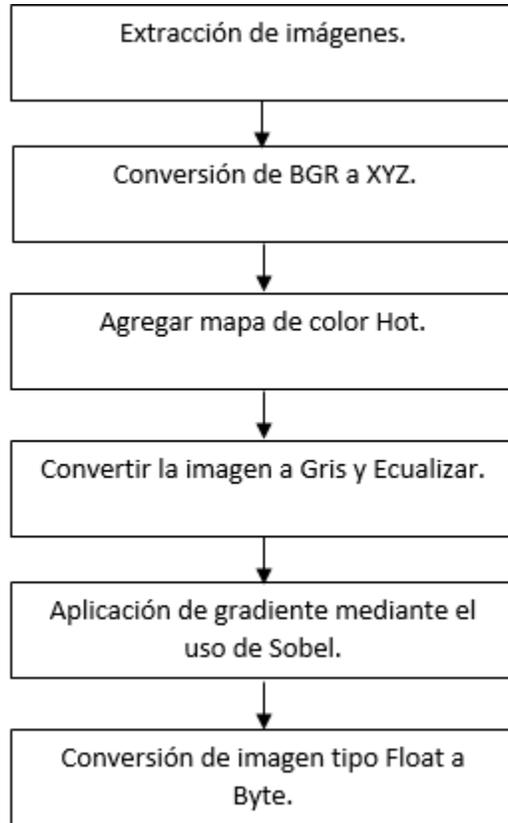


Figura 26. Flujo de proceso combinar2.

3.3.2.3 MORFOLOGÍA:

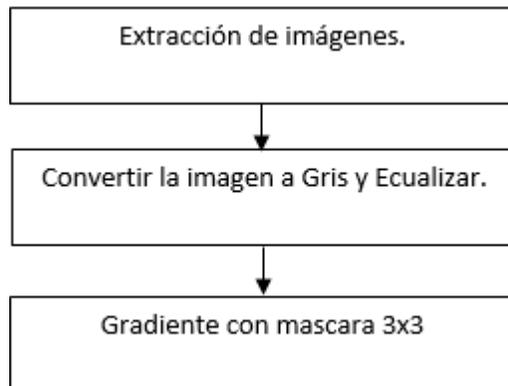


Figura 37. Flujo de proceso Morfología.

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

3.3.2.4 COLOR:

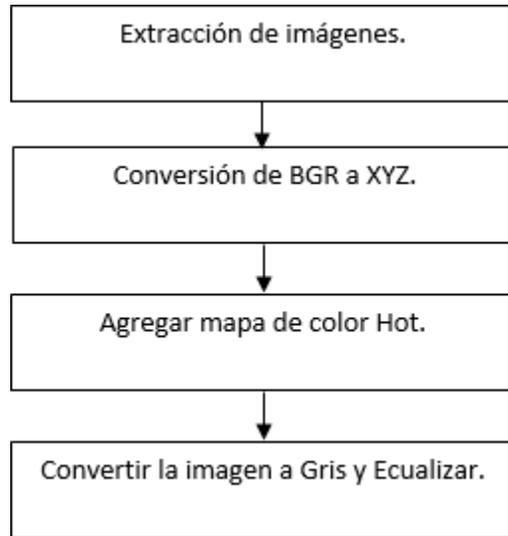


Figura 38. Flujo de proceso color.

3.3.2.5 ECUALIZACIÓN:

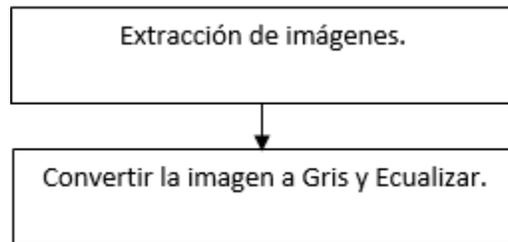


Figura 39. Flujo de proceso Ecuilización.

3.4 RECONOCIMIENTO FACIAL CON COMPUTACIÓN COGNITIVA

La metodología de la computación cognitiva se basa en el uso de redes neuronales profundas, convencionales o recurrentes y principalmente en el uso del aprendizaje profundo; que es responsable de inferir abstracciones de nivel superior a partir de la entrega de datos, utilizando un procesamiento que se basa en transformaciones lineales como no lineales, estas transformaciones son aquellas que se aplican de forma repetitiva a las capas de las redes neuronales mencionadas anteriormente.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

En el desarrollo de este proyecto el método elegido para aplicar la computación cognitiva fue mediante el uso del método Rekognition, el cual trabaja con redes neurales profundas que ya vienen pre-entrenadas en varias capas; sin embargo, tales redes neuronales ponen a disposición algunas capas que pueden ser utilizadas en el desarrollo del reconocimiento facial. Esto se hace comparando dos caras, la imagen de entrada y la imagen de comparación que es parte del conjunto o conjunto de imágenes; este conjunto se forma con la base de datos de personas, que se compone de una imagen por persona, tal como se explica en el desarrollo de la metodología. La comparación de estas dos caras se realiza a través de la medición de similitudes que se verifican con cada imagen dentro del conjunto; que analiza los atributos faciales e identifica los datos demográficos gracias a las funciones proporcionadas por el método; entre ellos se encuentran la comparación, el reconocimiento y la predicción que, junto con la ayuda de parámetros como los umbrales de confianza, permiten un reconociendo más riguroso reduciendo así los falsos positivos. Para la implementación de computación cognitiva inicialmente se debe crear una cuenta, ya que esta se necesita para poder usar el api y la consola. Una vez realizado esto se procede a crear un bucket el cual va a contener las imágenes y los labels, El bucket es proporcionado por el servicio, de manera que se puede subir una colección de rostros, cada uno de ellos con su nombre o label que permitirá su posterior identificación y reconocimiento. Para ello se utiliza el comando proporcionado por el api para la detección de rostros, la cual trabaja en imágenes de extensión jpeg o png. En esta implementación solo se cargó una sola imagen por persona, es decir que de las ocho personas que fueron cargadas en el bucket se montó una sola imagen con la cual se realizó el ejercicio y se obtuvieron los resultados. Una vez identificado el rostro se procede a una comparación, para esto se utiliza un dataset, el cual fue el bucket ya creado anteriormente y se utiliza una imagen de entrada con la cual se va a realizar el match. Una vez finalizado todo el proceso realizado por el servicio, este nos arroja como resultado final un json, el cual puede contener varios vectores dependiendo si se detectó y se comparó con varios rostros.

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

```
[
  {
    "Face": {
      "BoundingBox": {
        "Height": 1.08679,
        "Left": -0.0493294,
        "Top": -0.070911,
        "Width": 1.08679
      },
      "Confidence": 99.9625,
      "ExternalImageId": "JUAN",
      "FaceId": "83e83f76-0487-5200-ba73-cccb562834b7",
      "ImageId": "5919e942-cb30-5169-b207-3ea82e99d3d8"
    },
    "Similarity": 100
  }
]
```

Figura 40: Ejemplo de json

Dicho json proporciona el margen o boundary box indicando las características del rostro, el label o nombre de la imagen detectada y además proporciona la similitud con el rostro o con los rostros detectados; dicha similitud entre más cerca al valor uno más parecido entre las imágenes. Cabe decir que el servicio solo necesita un dataset el cual va a ser la fuente o bucket y luego necesita la imagen de entrada para hacer la detección, reconocimiento y comparación.

Hay que tener en cuenta que se programó para que solo generara una clasificación, si la imagen que se desea evaluar o comparar supera un porcentaje de acierto mayor al 90%; de presentarse el caso contrario el json devuelve [] que significa vacío.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

4. RESULTADOS Y DISCUSIÓN

Para todas las metodologías, es decir, las máquinas vectoriales de soporte, se evaluaron el clasificador en cascada y la detección de puntos con ocho personas. Se basó en un lapso de tiempo de aproximadamente dos minutos frente a la cámara web para que cada persona permaneciera frente a la cámara hasta que la metodología evaluara 120 frames o generara 120 calificaciones. El tiempo de dos minutos se definió como un estándar ya que cada técnica generó una clasificación por segundo, es decir, genera una evaluación de 60 frames por minuto. Hay que tener en cuenta que los resultados mencionados en este proyecto se encontraban en un entorno controlado, además de la posición de la cara con respecto a la cámara es frontal.

El criterio que se le dio a cada persona por su clasificación era de un rango del cual dependiendo del número de clasificaciones positivas contra las negativas se le asignó un valor que dependía del número de éxitos.

A continuación, se muestran los resultados de cada metodología en donde se identifica cual versión de cada técnica brindo mejores resultados por medio de graficas de barras.

4.1 MÁQUINAS DE VECTORES DE SOPORTE (SVM)

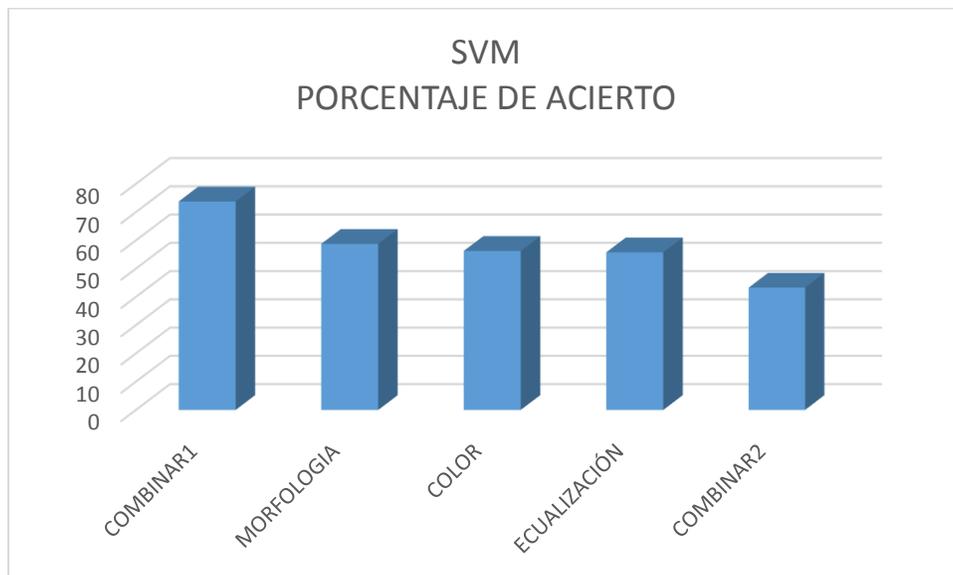
Para SVM como se mencionó en la metodología de este proyecto, se generaron varias versiones en donde se realizaron diferentes tratamientos de la imagen para comparar los resultados y así determinar cuál es la mejor versión para dicho método.

En la siguiente tabla y gráfica organizadas de forma descendente se podrá apreciar el porcentaje de aciertos de las muestras tomadas.

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

SVM	
MÉTODO	PORCENTAJE DE ACIERTO
COMBINAR1	73,75
MORFOLOGIA	58,75
COLOR	56,25
ECUALIZACIÓN	55,75
COMBINAR2	43,25

Tabla 1. Resultados SVM.



Grafica 1. Resultados SVM.

4.2 CASCADA

Para el método cascada se usaron diferentes modelos y diferentes versiones como en el método anterior. Las clases con las que se generaron las pruebas fueron:

MODELO	ID
FisherFaceRecognizer	F
EigenFaceRecognizer	E
LBPHFaceRecognizer	L

Tabla 2. Clases de Cascada

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

En la tabla 2 se le agrega un identificador como se muestra en la columna ID (F, E, L), estos identificadores se utilizan como prefijo en los resultados que se aprecian a continuación.

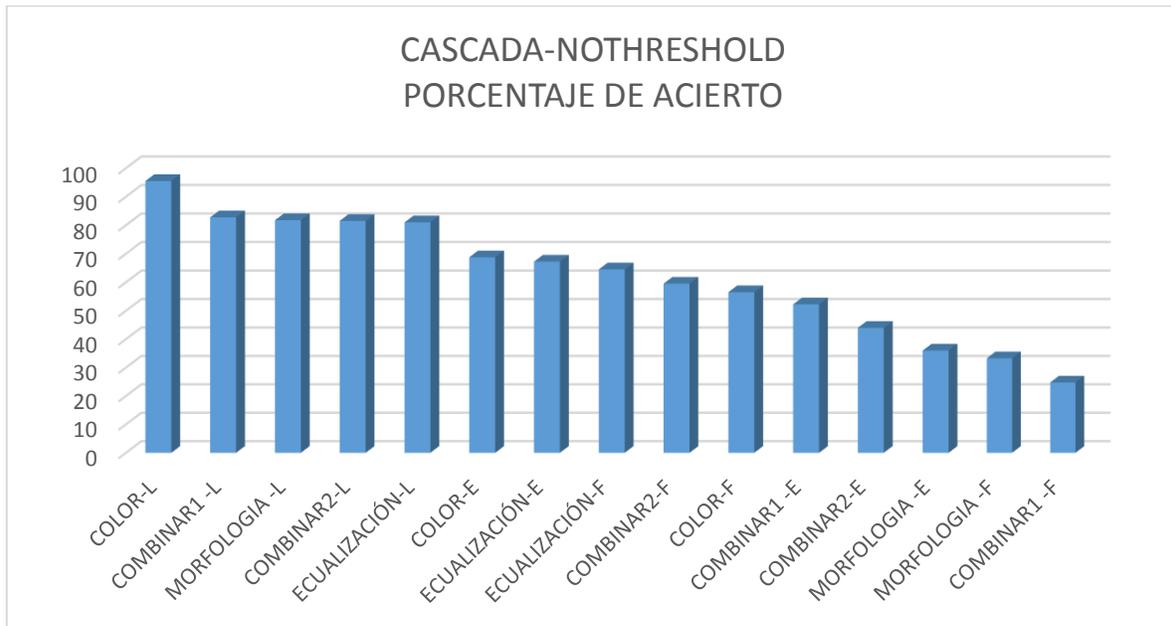
Además, se realizó pruebas por dos caminos distintos en el clasificador de cascada, es decir se hicieron pruebas con un umbral (threshold) como sin umbral (nothreshold); esto para obtener diferentes resultados y generar una comparación más exhaustiva.

En las siguientes tablas y gráficas organizadas de forma descendente se puede apreciar el porcentaje de aciertos de las muestras tomadas tanto para cascada nothreshold como para cascada threshold.

CASCADA-NOTHRESHOLD	
MÉTODO	PORCENTAJE DE ACIERTO
COLOR-L	95,5
COMBINAR1 -L	82,75
MORFOLOGIA -L	81,75
COMBINAR2-L	81,5
ECUALIZACIÓN-L	81
COLOR-E	68,75
ECUALIZACIÓN-E	67,25
ECUALIZACIÓN-F	64,5
COMBINAR2-F	59,5
COLOR-F	56,5
COMBINAR1 -E	52,25
COMBINAR2-E	44
MORFOLOGIA -E	36
MORFOLOGIA -F	33,25
COMBINAR1 -F	24,75

Tabla 3. Resultados Cascada - NOTHRESHOLD

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

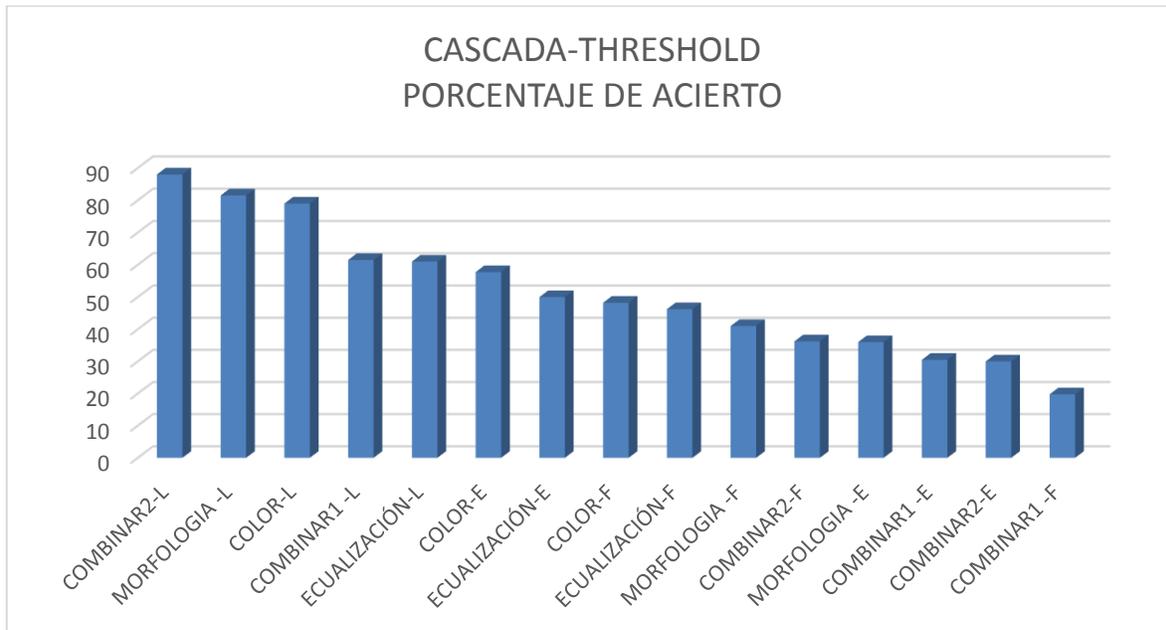


Grafica 2. Resultados Cascada - NOTHRESHOLD

CASCADA-THRESHOLD	
MÉTODO	PORCENTAJE DE ACIERTO
COMBINAR2-L	88
MORFOLOGIA -L	81,5
COLOR-L	79
COMBINAR1 -L	61,5
ECUALIZACIÓN-L	61
COLOR-E	57,75
ECUALIZACIÓN-E	50
COLOR-F	48,25
ECUALIZACIÓN-F	46,25
MORFOLOGIA -F	41
COMBINAR2-F	36,25
MORFOLOGIA -E	36
COMBINAR1 -E	30,5
COMBINAR2-E	30
COMBINAR1 -F	19,75

Tabla 4. Resultados Cascada - THRESHOLD

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22



Graficas 3. Resultados Cascada – THRESHOLD

4.3 DETECCIÓN POR PUNTOS

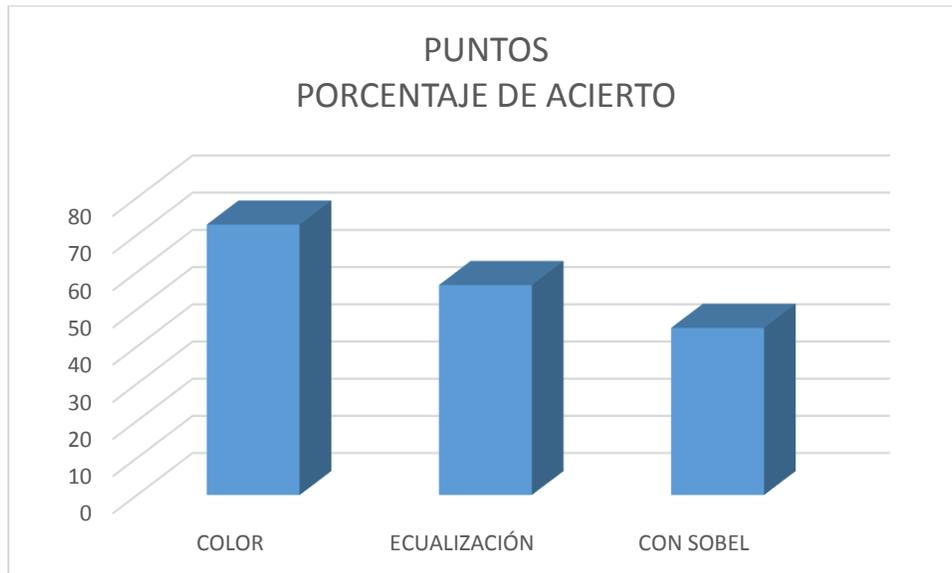
Para el método en puntos se generaron diferentes versiones con diferentes tratamientos de imagen como en los demás métodos. Aunque se modificaron algunos criterios de evaluación; en este método se realizaron 50 clasificaciones para evaluar el porcentaje de acierto por persona.

En la siguiente tabla y gráfica organizadas de forma descendente se podrá apreciar el porcentaje de acierto de las muestras tomadas para el método de detección por puntos.

MÉTODO	PORCENTAJE DE ACIERTO
COLOR	72,75
ECUALIZACIÓN	56,5
CON SOBEL	45

Tabla 5. Resultados Detección por puntos

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22



Grafica 4. Resultados Detección por puntos

4.4 RESUMEN RESULTADOS DE COMPARACIÓN

Finalmente se determinó el mejor método por medio de la matriz de confusión.

	1	2	3	4	5	6	7	8
1	96	0	0	0	10	5	9	0
2	0	91	11	6	0	0	0	12
3	0	10	101	5	0	0	0	4
4	0	4	4	108	0	0	0	4
5	22	0	10	0	48	20	20	0
6	9	0	0	0	0	96	15	0
7	13	0	0	0	5	6	96	0
8	0	16	18	14	0	0	0	72

Figura 41. Matriz de Confusión SVM.

$$\text{Precisión} = \text{TP} / (\text{TP} + \text{FP})$$

$$\text{SVM} = 708 / (708 + 252) = 0.7375 = 73\%$$

De igual manera haciendo la sumatoria de la precisión por etiqueta da el mismo resultado de 73%

$$\text{Es decir } (0.8 + 0.8 + 0.8 + 0.9 + 0.4 + 0.8 + 0.8 + 0.6) / 8 = 0.7375 = 73\%$$

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

	1	2	3	4	5	6	7	8
1	118	0	0	0	0	0	2	0
2	0	118	1	1	0	0	0	0
3	0	3	113	4	0	0	0	0
4	0	0	1	118	0	0	0	1
5	5	0	0	0	108	2	5	0
6	1	0	0	0	0	118	1	0
7	2	0	0	0	0	0	118	0
8	0	5	2	5	0	0	0	108

Figura 42. Matriz de Confusión Cascada.

$$\text{Cascada} = 919 / (919 + 41) = 0.957 = 95.7\%$$

$$(0.98 + 0.98 + 0.94 + 0.98 + 0.9 + 0.98 + 0.98 + 0.9) / 8 = 0.955 = 95.5\%$$

	1	2	3	4	5	6	7	8
1	96	0	0	0	7	4	13	0
2	0	24	28	40	0	0	0	28
3	0	29	48	28	0	0	0	15
4	0	4	3	113	0	0	0	0
5	1	0	0	0	115	2	2	0
6	19	0	0	0	0	84	17	0
7	6	0	0	0	0	6	108	0
8	0	5	3	2	0	0	0	110

Figura 43. Matriz de Confusión Puntos.

$$\text{Puntos} = 698 / (698 + 262) = 0.727 = 72.7\%$$

$$(0.8 + 0.2 + 0.4 + 0.94 + 0.95 + 0.7 + 0.9 + 0.91) / 8 = 0.725 = 72.5\%$$

Luego de los resultados vistos anteriormente las conclusiones son las siguientes:

El método en cascada con el modelo LBPHFaceRecognizer (nothreshold) utilizando la versión de color fue la que entrego mejor resultado con respecto a su precisión en un 95,5%.

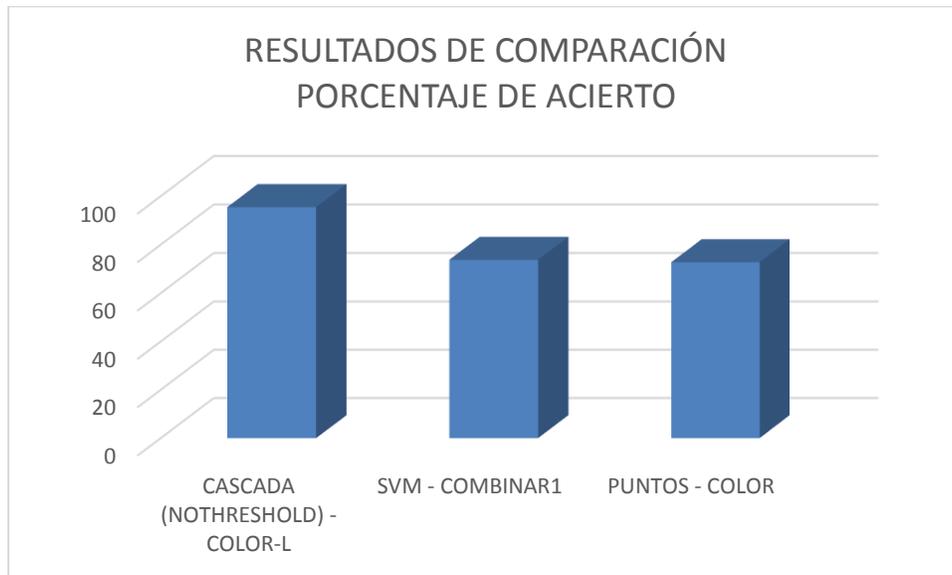
El método que lo sigue fue el de máquinas de vectores de soporte con la versión combinar 1 con un 73,75% y en el tercer puesto del comparativo quedo el método de detección por puntos con la versión de color con un porcentaje del 72,75%.

En la siguiente tabla y grafica se puede ver reflejado el comparativo entre los métodos:

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

RESULTADO DE COMPARACIÓN	
MÉTODO	PRECISIÓN
CASCADA(NOTHRESHOLD) - COLOR- L	95,7
SVM - COMBINAR1	73,75
PUNTOS - COLOR	72,7

Tabla 6. Resultado de Comparativo



Gráfica 5. Resultado de Comparativo

4.5 COMPARACIÓN DE TIEMPOS

A continuación, se enumerarán los tiempos que se plantearon en los objetivos, los cuales se midieron en cada uno de los tres métodos de reconocimiento facial planteados en el presente proyecto.

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

4.5.1 REENTRENAMIENTO – ENTRENAMIENTO

MÉTODO/VERSIÓN	SVM	CASCADA	PUNTOS
ECUALIZACIÓN	~6-8 minutos	~20-22 minutos	~1,66-3,33 minutos
COMBINAR1	~13-14 minutos	~20-22 minutos	~1,66-3,33 minutos
COMBINAR2	~11-12 minutos	~20-22 minutos	~1,66-3,33 minutos
MORFOLOGIA	9-11 minutos	~20-22 minutos	~1,66-3,33 minutos
COLOR	~7-8 minutos	~20-22 minutos	~1,66-3,33 minutos

Tabla 7. Resultados de Entrenamiento

SVM no tiene reentrenamiento en opencv, tiene que cambiarse todo el dataset, lo cual significa generar una nueva extracción y un nuevo archivo XML, eso daría un tiempo de extracción de aproximadamente ~100-200 segundos más el tiempo de generar el XML que es el que se muestra en la tabla 7. También se debe tener en cuenta que dependiendo del tipo de proceso este puede incrementar el tiempo, por ejemplo, para color sería ~100-200 segundos más ~7-8 minutos lo cual daría algo como ~ 8-12 minutos.

Además, hay que tener en cuenta que durante el proyecto se desarrolló el entrenamiento bajo las funciones del open cv, en este caso para generar el archivo XML solo se usó la función linear por motivos de que la función inter desgastaba más tiempo. Se adjunta una tabla con el tiempo de generación del archivo XML.

VERSIÓN	TIEMPO
ECUALIZACIÓN	~2:45:00
COMBINAR1	~4:00:00
COMBINAR2	~4:00:00
MORFOLOGIA	~4:00:00
COLOR	~3:00:00

Tabla 8. Tiempos entrenamiento SVM con la función inter.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

- El reentrenamiento en cascada necesita cambiar por completo el dataset y el tiempo de reentrenamiento usa la misma fórmula que SVM, es decir el tiempo de extracción más el tiempo de generación del archivo.

Además, hay que tener en cuenta como se vio en la tabla 7, que cascada tiene un tiempo de entre 20 minutos a 22 minutos, pero cabe recordar que se hicieron pruebas con un umbral (trheshold) como sin umbral (no trheshold) y esto multiplicaría los tiempos para generar los archivos siendo un tiempo aproximado de 40 a 44 minutos.

- La detección por puntos no requiere entrenamiento ya que se basa en la base de datos de rostros que se tenga, por lo cual el tiempo de entrenamiento es igual al reentrenamiento; para el método de detección por puntos tanto con homografía como sin homografía se demora ~100-200 segundos tanto para entrenamiento como para reentrenamiento, esto puede variar dependiendo de la cantidad de imágenes, este proyecto se implementó con 200 imágenes.

- Los tiempos obtenidos en este análisis de resultados pueden variar dependiendo del procesador y de las características del computador.

A continuación, se listan las características del computador con las que se hicieron todas las pruebas y se generaron los archivos.

COMPUTADOR A

- **Sistema operativo:** Windows 8.1.
- **Procesador:** intel(r) core(tm) i5-4200u CPU @ 1.60ghz 2.30ghz.
- **RAM:** 8,00 GB (7,89 GB utilizable).

Hay que tener en cuenta que la implementación hecha de cada metodología se realizó adicionalmente en dos computadores más, para analizar su comportamiento con diferentes características y así evidenciar su respuesta de cada una. Las características fueron:

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

COMPUTADOR B

- **Sistema operativo:** Windows 10.
- **Procesador:** Intel(R) Core(TM) i7-4510U CPU @ 2.00GHz 2601 MHz, 2 procesadores principales, 4 procesadores lógicos.
- **RAM:** 8,00 GB.

COMPUTADOR C

- **Sistema operativo:** Windows 7.
- **Procesador:** Intel(r) core(tm) i3-3110m CPU @ 2.40ghz.
- **RAM:** 4,00 GB (3,88 GB utilizable).

4.6 RECONOCIMIENTO FACIAL COGNITIVO

En la siguiente sección se muestra los resultados de utilizar computación cognitiva para el reconocimiento facial. Por costos en el consumo del servicio, se utilizaron menos frames para la toma de resultados.

Para la recolección de los datos se utilizaron 20 frames para las mismas 8 personas que en los métodos convencionales. La clasificación entregada por el servicio es un archivo json, en los cuales se abstrae el porcentaje de similitud de la clasificación del frame para luego realizar la sumatoria de los 20 frames por persona y calcular un promedio de acierto.

Los resultados entregaron un porcentaje de acierto del 99.51720205%.

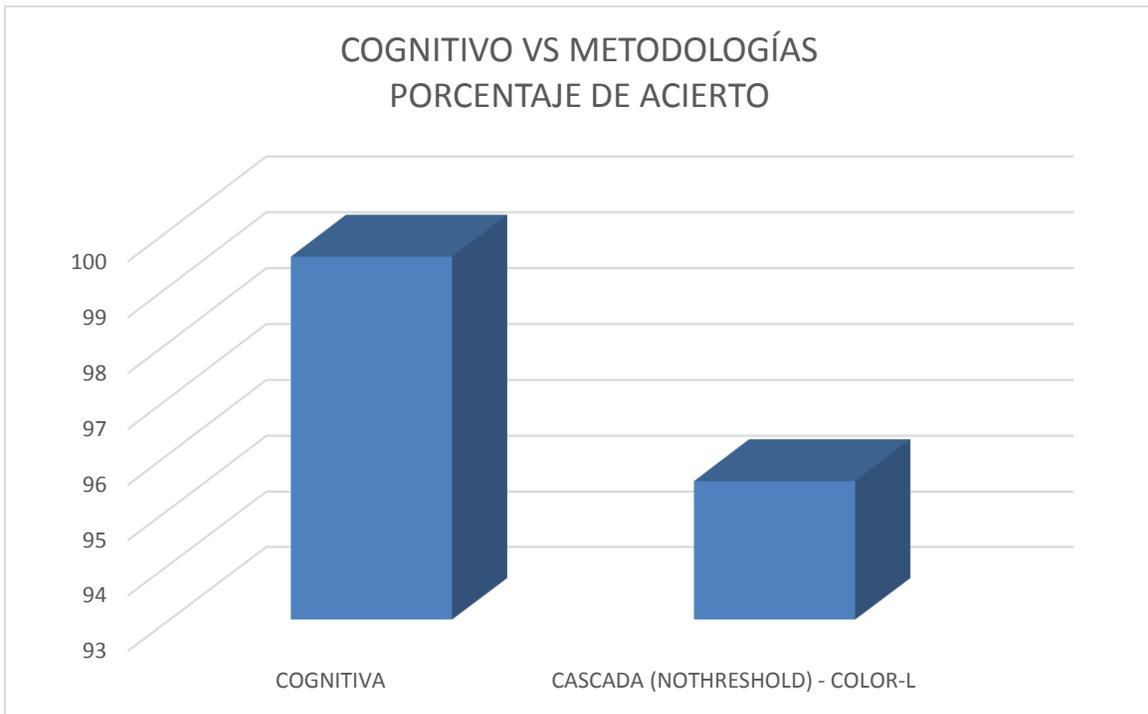
4.6.1 COMPUTACIÓN COGNITIVA vs METODOLOGÍA CON MEJOR RESULTADO

A continuación, se presentan una tabla y una gráfica en donde se muestran los resultados obtenidos con respecto al porcentaje de acierto entre la computación cognitiva y el método que entrego el mejor resultado, en este caso cascada (nothreshold) aplicando tratamiento de color con el modelo LBPHFaceRecognizer.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

COGNITIVO VS METODOLOGÍAS	
MÉTODO	PORCENTAJE DE ACIERTO
COGNITIVA	99,51720205
CASCADA (NOTHRESHOLD) - COLOR-L	95,5

Tabla 9. Comparación Cognitiva vs Cascada (nothreshold) – Color- L



Grafica 6. Comparación Cognitiva vs Cascada (nothreshold) – Color- L

Como se logra apreciar en la tabla 9 el porcentaje de acierto de ambos fue superior al 90 %. Pero la que entrego un mayor porcentaje de acierto fue la computación cognitiva, la cual fue de un 99.5172 teniendo casi una efectividad del 100%.

Con respecto a los tiempos de entrenamiento, la computación cognitiva tiene un tiempo aproximado a un segundo; esto es lo que se demora en cargar una imagen al bucket y asociarle un label. Caso contrario con el método de cascada que tiene un tiempo de entrenamiento entre 20 a 22 minutos aproximadamente.

El número de imágenes por persona para la implementación de la computación cognitiva

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

fue de una sola imagen siendo un total de ocho imágenes, que corresponden a las ocho personas con las que se obtuvieron y analizaron los resultados. Caso contrario con el número de imágenes que se utilizó para el método de cascada que fueron de 200 imágenes por persona siendo un total de 1600 imágenes.

Para realizar el reentrenamiento en la computación cognitiva solo se debe agregar la nueva imagen al bucket y asociarle un label; teniendo así el mismo tiempo de la carga inicial al bucket, el cual es de aproximadamente un segundo. Para el método en cascada, el reentrenamiento necesita cambiar el dataset por completo lo cual genera un tiempo estimado de 20 minutos.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

5. CONCLUSIONES, RECOMENDACIONES Y TRABAJO FUTURO

CONCLUSIONES

- Se realizó un análisis de los resultados obtenidos en las tres metodologías de svm, cascada y detección de puntos. En el que se concluyó que el método en cascada (no umbral) fue el que dio un mejor resultado con un porcentaje de acierto del 95,5%. El método seguido fue el de máquinas de soporte vectorial con 73,75% y en el tercer lugar de la comparación fue el método de detección de manchas con un porcentaje del 72,75%.
- Con la clasificación anterior, se generó la comparación entre el método de mejor desempeño y la computación cognitiva. Lo cual dejó como resultado de que la computación cognitiva fue superior con una efectividad del 99.5172. Además de que la computación cognitiva tiene flexibilidad en el reentrenamiento, posibilidad que no contempla el método en cascada.
- Los tiempos de clasificación en las tres metodologías según el análisis de resultados del presente proyecto entregó que los métodos de SVM y cascada generó una clasificación del rostro cada segundo por frame. Para la detección por puntos se tardó entre 15 a 20 segundos para evaluar cada frame (Resultados del computador A). Con respecto a los computadores B y C entregó para SVM y cascada una clasificación por segundo, no existió variación con respecto al computador principal del proyecto; aunque para la detección por puntos si se apreció un cambio, para el computador B fue entre 12 y 17 segundos y para el computador C el tiempo fue entre 23 y 28 segundos.
- El reentrenamiento no es posible para SVM y cascada con implementaciones hechas. Para la detección de puntos, si es posible, y los tiempos de re-entrenamiento

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

son los mismos que los tiempos de entrenamiento, que pueden variar entre 1,66 y 3,33 minutos.

- La comparación entre el método en cascada que dio el mejor resultado, con la implementación de la computación cognitiva, dicha comparación mostró la amplia ventaja que los sistemas cognitivos tienen sobre los métodos convencionales y pueden ser evidenciados o concluidos en el presente proyecto donde los resultados muestran un porcentaje de acierto superior al 99%, utilizando como insumo el entrenamiento de una imagen por persona, además de la respuesta en la detección y clasificación que no excede de un segundo; vale la pena señalar que el re-entrenamiento no afecta al conjunto de datos actual, sino que consiste en agregar una imagen asociada a una etiqueta; caso contrario con el método en cascada que cuenta con un re-entrenamiento, donde toca realizar nuevamente el entrenamiento por completo de todas las muestras, teniendo un tiempo superior a los 20 min. Aunque los resultados del método en cascada fueron superiores al 95%, tiene sus limitaciones a lo largo del proyecto.

RECOMENDACIONES

Durante el desarrollo del proyecto y durante la realización de varios métodos, primeramente se obtuvo en cuenta el tamaño completo de la imagen del rostro, pero a medida que se realizaron pruebas y por falta de un mejor equipo (hardware), el tamaño se redujo hasta el punto de 60x60, esto con el fin de que las estructuras de datos se pudieran almacenar; de manera que los resultados obtenidos fueron buenos en algunos métodos y aceptables en otros, además gracias a la reducción de tamaño, no se necesita tanto procesador. Cabe decir que, debido a lo anteriormente mencionado el tamaño de las imágenes de rostros pueden influir en la obtención de mejores resultados.

En el desarrollo de los modelos SVM, cascada y detección por puntos, se utilizaron diferentes métodos. Es decir que para el modelo SVM se utilizaron métodos de clasificación Linear, polinomial, inter, sigmoide y rbf; los cuales dieron buenos resultados, a excepción

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

de los últimos dos. De manera que es posible mejorar dichos resultados de estos métodos si se utilizan diferentes parámetros o si se realiza un procesamiento diferente en las imágenes. Con el método de cascada se utilizó un umbral y un número de componentes determinado, mas no se probaron otras combinaciones, esto deja la posibilidad de probar o intentar otros valores para dichos componentes y umbrales, que pueden mejorar los resultados que se obtuvieron en el trabajo. Finalmente en el modelo de detección de puntos se utilizó homografías y el gfttdetector, pero con las homografías no se obtuvieron los resultados esperados debido al criterio de evaluación que determinaba la mejor homografía, dicho criterio se puede cambiar o modificar para mejorar la evaluación de la imagen con respecto a su homografía, en cuanto al gfttdetector se utilizó los parámetros recomendados, debido a que durante el desarrollo de este método, al modificar los parámetros se obtuvieron malos resultados en comparación a los recomendados; de manera que estos parámetros si se pueden modificar pero las imágenes necesitan otro procesamiento. Finalmente cabe decir que el tamaño de las imágenes en detección por puntos era de 499 x 499, pero este se cambió a 60 x 60 por motivos de procesamiento computacional (hardware).

TRABAJOS A FUTUROS

Con el desarrollo y comparación de diferentes métodos, se deja abierta la posibilidad de mejorarlos en el sentido que, las imágenes no necesiten ningún procesamiento o que, al combinarlos con métodos de inteligencia artificial como redes neuronales, estos puedan mejorar en cuanto a la clasificación y a la modificación del dataset cuando se agrega una nueva persona. También se deja la posibilidad de probar diferentes criterios de evaluación y comparación en la detección por puntos para mejorar su porcentaje de acierto, cabe mencionar que estos modelos se desarrollaron con las funciones de apoyo de opencv; por lo cual se pueden probar otros tipos de clasificadores, como por ejemplo un nuevo svm que permita un crecimiento incremental, de manera tal que el re entrenamiento sea posible y no se necesite cambiar todo el dataset.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

Con respecto al procesamiento de imágenes y mejora de cada uno de los métodos desarrollados durante este proyecto, se deja la posibilidad de utilizar otros ambientes de desarrollo, es decir, que durante todo el proyecto se utilizó un ambiente controlado; pero sin tener en cuenta una luminosidad perfecta y corrección de blanco. Debido a lo anterior se pueden mejorar los resultados si se utilizan técnicas de fotografía o si se tiene más experiencia en dicho ámbito laboral, de esta manera las imágenes no necesitarían tanto procesamiento y mejoraría las velocidades de clasificación, reduciendo también las especificaciones de un procesador. Cabe decir que las imágenes se pueden separar utilizando solo un canal antes de convertirse a escala de gris, en vez de utilizar los tres canales, de manera que los resultados podrían mejorar con respecto al uso de color XYZ. Finalmente el presente proyecto se desarrolló enfatizado en comparar los porcentajes de aciertos de los métodos SVM, clasificador en cascadas y por puntos. Se propone aplicar un comparativo basado en otras métricas como: sensibilidad, especificidad, F-score. Las cuales podrán ampliar y brindar otras conclusiones para definir cuál metodología es la mejor.

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

REFERENCIAS

- Aliradi, R. (2013). *Detection of Facial Components Based on SVM Classification and Invariant Feature*. In *Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*, . IEEE/WIC/ACM International Joint Conferences on.
- Alonso, J., & Matos, M. (. (2008). *La computación cognitiva, el valor de unir máquinas y humanos*. Obtenido de harvard-deusto.com: <https://www.harvard-deusto.com/la-computacion-cognitiva-el-valor-de-unir-maquinas-y-humanos>
- Ballesta, A., Gil, O., & Reinoso, D. (2010). *Análisis de Detectores y Descriptores de Características Visuales en SLAM en Ambientes Internos y Externos*. Alicante,: Universidad Miguel Hernández.
- Bradski, G., & Kaehler, A. (2008). *“Learning OpenCV: Computer Vision with the OpenCV Library”*. . O’Reilly Media.
- Cevikalp, H., Triggs, B., & Franc, V. (2013). *Face and landmark detection by using cascade of classifiers*. In *Automatic Face and Gesture Recognition (FG)*. IEEE International Conference and Workshops on.
- Coca, L. F. (2008). *Implementación De Filtros Morfológicos Utilizados En El Procesamiento De Imágenes Digitales En Un Dispositivo Lógico Programable*. Obtenido de scielo.org.ve: http://www.scielo.org.ve/scielo.php?script=sci_arttext&pid=S1316-
- Cohen, I., Sebe, N., Garg, A., Chen, L. S., & Huang, T. S. (2003). *Facial expression recognition from video sequences: temporal and static modeling*. *Computer Vision and image understanding*.
- Guerrero, W. (2014). *Detección y descripción de puntos característicos en imágenes multispectrales utilizando esquemas clásicos*. Revista Tecnológica.
- Hussain, A. (2013). *Cognitive computation: A case study in cognitive control of autonomous systems and some future directions*. In *Neural Networks (IJCNN)*. International Joint Conference on. .

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

Jian-zheng, L. (2011). *Fully automatic and quickly facial feature point detection based on LK algorithm. In Networked Computing and Advanced Information Management (NCM)*. International Conference on.

Kelly, J. (2015). *Computing, cognition and the future of knowing. How humans and machines are forging a new age of understanding*. IBM Research and Solutions Portfolio.

Monadjemi, T. (2011). *Face Detection and Recognition Using Skin Color and AdaBoost Algorithm Combined with Gabor Features and SVM Classifier. In Multimedia and Signal Processing (CMSP)*. International Conference on.

Panning, A., Al-Hamadi, A., & Michaelis, B. (2012). *Facial feature point detection using simplified gabor wavelets and confidence-based grouping*. IEEE International Conference on Systems, Man, and Cybernetics (SMC).

Patil, R., Sahula, V., & Mandal, A. (2011). *Automatic detection of facial feature points in image sequences. In Image Information Processing (ICIIP)*. International Conference on.

Rong, W., Chen, H., Liu, J., & Haeusler, R. (2009). *Mosaicing of microscope images based on surf. In Image and Vision Computing New Zealand*. New Zealand: International Conference.

Sariyanidi, E. (2012). *Real time face detection and recognition*. Signal Processing and Communications Applications Conference (SIU).

Xue, M., & Zhu, C. (2009). *A study and application on machine learning of artificial intelligence. In Artificial Intelligence*. International Joint Conference on .

Zhang, P. (2008). *A video-based face detection and recognition system using cascade face verification modules*. Applied Imagery Pattern Recognition Workshop.

