



**BENEMÉRITA UNIVERSIDAD  
AUTÓNOMA DE PUEBLA**  
FACULTAD DE CIENCIAS DE LA COMPUTACIÓN

---

**Clasificador de edad en imágenes digitales  
basado en métodos estadísticos.**

TESIS PRESENTADA PARA OBTENER EL TÍTULO DE:

**LICENCIADO EN INGENIERÍA EN CIENCIAS DE LA  
COMPUTACIÓN**

PRESENTA:

**ROBERTO GUILLERMO TORRES LÓPEZ**

ASESORES:

Dr. LUIS ENRIQUE COLMENARES GUILLEN

Dra. MAYA CARRILLO RUIZ

PUEBLA, PUE.

DICIEMBRE 2017

## **Agradecimientos**

Mi primer agradecimiento, es para mí asesor de tesis, Luis Enrique Colmenares ya que sin sus enseñanzas, su apoyo y su absoluta confianza en mi capacidad investigadora, esta tesis nunca habría sido posible.

Tampoco habría sido posible sin el apoyo de mis padres, por darme la oportunidad de realizar la carrera y haberme aguantado y guiado todos estos años. También agradecer a mis familiares por el apoyo e interés constante.

Por último y no menos importante, quiero agradecer a todos a aquellas personas que conocí y me acompañaron durante los años de la carrera, con quien disfrute, sufrí, viví y sobreviví los días y las noches de proyectos.

Y finalmente, gracias a ti querido lector. Espero que disfrutes leyendo esta humilde tesis siquiera una pequeña parte de lo que yo he disfrutado escribiéndola.

Roberto Torres.

Puebla Puebla, Diciembre 2017

## Tabla de contenido

<b>CAPÍTULO 1.....</b>	<b>8</b>
INTRODUCCIÓN .....	9
<i>Control de contenidos para menores de edad.....</i>	<i>9</i>
<i>Publicidad segmentada .....</i>	<i>10</i>
<i>Detector de pedofilia.....</i>	<i>11</i>
MOTIVACIÓN.....	12
JUSTIFICACIÓN .....	12
HIPÓTESIS .....	13
<i>Hipótesis General.....</i>	<i>13</i>
<i>Hipótesis Específicas.....</i>	<i>13</i>
OBJETIVOS .....	13
<i>Objetivo General.....</i>	<i>13</i>
<i>Objetivos Específicos .....</i>	<i>14</i>
RESUMEN CAPÍTULO 1 .....	14
<b>CAPÍTULO 2.....</b>	<b>15</b>
MARCO CONCEPTUAL.....	15
<i>Envejecimiento .....</i>	<i>15</i>
<i>Estimación de edad .....</i>	<i>16</i>
MARCO TEÓRICO.....	17
<i>Modelos de representación de características.....</i>	<i>17</i>
Modelos Antropométricos.....	18

Apariencia Activa.....	20
Patrones de envejecimiento.....	21
Manifold de edad.....	23
Apariencia .....	24
Fisherfaces como representación de características .....	26
<i>Técnicas de estimación de edad.....</i>	<i>31</i>
Clasificación por vecindad.....	31
Máquina de soporte vectorial .....	32
Redes neuronales.....	34
Modelos de regresión .....	36
Regresión de soporte vectorial .....	37
<i>Biblioteca de procesamiento OpenCV .....</i>	<i>37</i>
RESUMEN CAPÍTULO 2 .....	38
<b>CAPÍTULO 3.....</b>	<b>39</b>
METODOLOGÍA.....	39
<i>Bases de datos de envejecimiento .....</i>	<i>40</i>
IMDB-Wiki data set.....	40
FG-Net.....	41
<i>Detección del rostro .....</i>	<i>41</i>
<i>Ruido y tipos de ruidos.....</i>	<i>44</i>
<i>Filtrado de las imágenes .....</i>	<i>45</i>
<i>Propuesta del clasificador .....</i>	<i>48</i>

<i>Creación de los modelos</i> .....	49
<i>Clasificación de una nueva imagen</i> .....	50
RESUMEN CAPITULO 3 .....	51
<b>CAPÍTULO 4</b> .....	<b>52</b>
PRUEBAS .....	52
RESULTADOS.....	52
<i>Clasificación de mayoría y minoría de edad.</i> .....	52
<i>Estimación de edad.</i> .....	57
RESUMEN CAPITULO 4.....	60
<b>CONCLUSIONES</b> .....	<b>61</b>
<b>REFERENCIAS</b> .....	<b>63</b>
<b>APÉNDICE A COMPARACIÓN DE FILTROS</b> .....	<b>68</b>

## Lista de figuras

Figura 1. Efectos de la cirugía estética (Troell, 2014). .....	16
Figura 2. Efecto del maquillaje (Begoña, 2013). .....	16
Figura 3. Esquemas de la edad de un rostro desde los 30 a los 80 años (Fu Guo & Huang, 2010). 15	
Figura 4. Modelo antropométrico (Kwon & Lobo 1994,1999).....	19
Figura 5. Modelo antropométrico (Ramanathan y Chellappa 2006).....	19
Figura 6. Modelos de apariencia activa (Alejandro Torrisi, 2015). .....	21
Figura 7. Patrones de envejecimiento (Xin Geng, 2007). .....	22
Figura 8. Manifold de edad (Guodong Guo, 2007).....	24
Figura 9. Modelos de apariencia [imagen] disponible en: <a href="http://www.adelantandoelmundo.com/2015/04/averiguar-edad-persona.html">http://www.adelantandoelmundo.com/2015/04/averiguar-edad-persona.html</a> [recuperado 08 sep 2017].....	25
Figura 10. Grafico de valores vectoriales para imágenes de ancianos y jóvenes (Patterson 2014). 28	
Figura 11. Ejemplo de una gráfica ideal creada por el análisis discriminante lineal [imagen] disponible en: <a href="http://www.dtreg.com/LdaSepGood.jpg">http://www.dtreg.com/LdaSepGood.jpg</a> [recuperado 2017-09-20].....	30
Figura 12. Ejemplo de salida de Fisherface detallando las características de la clase adultos [imagen] disponible en: <a href="http://docs.opencv.org/trunk/_images/fisherface_0.png">http://docs.opencv.org/trunk/_images/fisherface_0.png</a> [recuperado 2017- 09-20]. .....	30
Figura 13. Máquinas de soporte vectorial (Canales, 2009).....	32
Figura 14. Redes neuronales (González-Rodríguez, A. F. 2014). .....	35

Figura 15. Redes neuronales. (2017). [imagen] disponible en: <a href="http://www.rsipvision.com/exploring-deep-learning/">http://www.rsipvision.com/exploring-deep-learning/</a> [recuperado 8 Sep. 2017].....	36
Figura 16. Modelo general del clasificador.....	39
Figura 17. Base de imágenes IMBD-Wiki.....	41
Figura 18. Ejemplo base de datos Fg-Net. ....	41
Figura 19. Detección y recorte del rostro.....	44
Figura 20. Imagen con ruido gaussiano. ....	44
Figura 21. Imagen con ruido sal y pimienta.....	45
Figura 22. Conversión de una imagen RGB a escala de grises.....	46
Figura 23. Ecualizado del histograma de una imagen.....	47
Figura 24. Preprocesado de una imagen.....	47
Figura 25. Módulos del clasificador.....	48
Figura 26. Creación de modelos Fisherfaces. ....	49
Figura 27. Fisherface mayores y menores de edad. ....	50
Figura 28 Proceso de clasificación de una nueva imagen.....	50
Imagen 29. Matriz de confusión.....	54
Figura 30. Interpretación de valores en la matriz de confusión. ....	54
Figura 31. Clasificación visual de mayoría y minoría de edad. ....	57
Figura 32. Estimación visual de edad.....	60

## Lista de tablas

Tabla 1. Características de la computadora de pruebas. ....	52
Tabla 2. Resultados con la base de datos Fg-Net. ....	56
Tabla 3. Comparación con trabajos relacionados en clasificación de mayoría y minoría de edad. .	57
Tabla 4. Resultados estimación de edad. ....	58
Tabla 5 Resultados con trabajos similares en estimación de edad. ....	58

# Divulgación científica

## Congresos:

- Colmenares Luis, Torres Roberto, Carrillo Maya, Albores Francisco. *Clasificador de edad en imágenes digitales usando métodos estadísticos*. 9º Congreso Mexicano de Inteligencia Artificial (COMIA 22 al 27 de mayo del 2017).
- Torres Roberto, Colmenares Luis, Carrillo Maya, Hernández José, Chavira Elsa. *Una Aproximación de un clasificador de edad en imágenes digitales usando métodos estadísticos como modelos de representación*. Congreso Nacional de Ciencias de la Computación (CONACIC 6 al 10 de noviembre del 2017).

# Capítulo 1

## Introducción

El desarrollo tecnológico iniciado con el Internet y sus aplicaciones muestra que la necesidad de identificar la edad de las personas que acceden a distintas páginas y servicios disponibles en la red se vuelve cada vez más importante, debido a que cada día aparecen en la red nuevas y diversas amenazas, en especial para los menores de edad, tales como “*Cyberbullying*” (acoso a través de Internet) o “*Grooming*” (suplantación de identidad por parte de un adulto que se hacen pasar por menor).

En este entorno se desenvuelven, relacionan y conviven virtualmente los niños y adolescentes que, al ser nativos digitales, cuentan con las habilidades propias del autoaprendizaje, que en muchos casos ignoran las precauciones necesarias para evitar los riesgos a los que están expuestos [1].

## Control de contenidos para menores de edad

Cuando un menor de edad navega libremente por Internet corre el riesgo de encontrar contenidos relacionados con apuestas, escenas de violencia, consumo de alcohol y drogas, además puede comunicarse con personas desconocidas que puedan engañar, seducir, abusar e inclusive desarrollar acciones ilícitas contra él, solicitando información personal como nombre, dirección, teléfono, aficiones o datos familiares.

Además un menor de edad puede acceder a páginas con material sexual explícito [2], debido a que existe una gran cantidad de pornografía en la red, donde una simple consulta a través de los motores de búsqueda, como Google, Yahoo o Bing, deja en evidencia que este tipo de sitios se cuentan por cientos de miles, por lo que basta con hacer un clic o con teclear la dirección

electrónica para acceder al contenido solicitado, caso contrario a la televisión donde existen restricciones de horarios o las revistas con material pornográfico en las cuales se puede limitar su venta exclusiva a los mayores de edad [3].

## **Publicidad segmentada**

Resulta relevante el comportamiento y la forma de consumo por parte de los infantes los cuales son receptivos ante todo lo que los rodea siendo un punto importante el que se destaca en [4] el cual dice que con solo 12 meses los niños son capaces de decodificar a la perfección un video de 20 segundos de duración, imitando el comportamiento y las reacciones de sus protagonistas. Este descubrimiento advierte el poder de la publicidad sobre las mentes de los menores de edad y se maximiza ya que desde el momento en que nace un niño está rodeado de anuncios publicitarios y marcas que representan empresas.

La inocencia de los niños es el factor fundamental que les impide ser objetivos a los anuncios publicitarios y al recibir bombardeos de información, el niño puede desarrollar formas de actuar inadecuadas para su desarrollo como ser humano.

Ante esto, el crecimiento y aceptación de Smart TV con cámara web integrada dentro de los hogares y la capacidad de las mismas para aprender sobre los contenidos que observa cada miembro de la familia [5] se deben implementar medidas para segmentar la publicidad por grupos etarios restringiendo los anuncios que hagan referencia a bebidas alcohólicas, cigarros o algún comportamiento apto para personas con mayor madurez mental.

## **Detector de pedofilia**

La pedofilia es una actividad que se ha incrementado con el desarrollo del Internet, una de las razones es el anonimato y la facilidad de hacer contacto tanto con menores de edad como con otros pedófilos. Varios proyectos e iniciativas para investigar dicho fenómeno se han desarrollado a nivel mundial como [6] y [7] donde agencias gubernamentales encargadas de minimizarlo utilizan primordialmente dos estrategias: páginas ficticias que hacen el papel de señuelo para detectar usuarios pedófilos, y agentes de policía que haciéndose pasar por niños logran engañarlos para su posterior arresto como se muestra en [8]. Se estima que el número de imágenes de abusos sexuales infantiles en Internet asciende a millones y el número de niños filmados o fotografiados es, probablemente de decenas de miles [9].

En la mayoría de imágenes de abusos sexuales infantiles aparecen niñas caucásicas, preadolescentes debido a que la mayor parte de las investigaciones realizadas hasta ahora se han llevado a cabo en países occidentales y que los delincuentes prefieren a menores que comparten sus propias etnias. También puede reflejar una mayor disponibilidad de las TIC así como de otras tecnologías para la grabación y distribución de imágenes en los países industrializados haciendo difícil calcular el número de sitios web en todo el mundo que muestran imágenes de abusos sexuales infantiles. En 2010 [11], la Internet Watch Foundation (IWF) identificó y tomó medidas contra aproximadamente 16.700 casos de abusos sexuales contra niños en distintas páginas web a nivel mundial en comparación con aproximadamente 10.600 dominios (URL) en páginas o sitios de Internet detectados en 2006 [12], el aumento puede atribuirse al cambio en las pautas de alojamiento de contenidos, el cual se envía a distintos sitios en vez de almacenar varias imágenes en una sola página web [13].

Ante esta problemática el laboratorio de análisis forense digital de la Benemérita Universidad Autónoma de Puebla desarrolla un sistema que a partir de una imagen con contenido pornográfico y un clasificador de edad pueda detectar los sitios que contengan o distribuyan pedofilia.

## **Motivación**

La clasificación automática de edad en imágenes faciales es un método biométrico con características no invasivas que puede ser utilizada sin el consentimiento y la colaboración del sujeto. Por ello brinda un vasto campo de aplicación en el mundo tecnológico actual, en especial, en lo que respecta a la seguridad informática de los menores de edad y publicidad segmentada.

Un prototipo funcional permitiría aplicar control de acceso, personalizar contenidos y detectar pedofilia en Internet.

## **Justificación**

En este trabajo de fin de carrera se busca una red segura para los menores de edad, mediante la implementación de mecanismos automáticos para proteger a los niños de personas potencialmente peligrosas con las que pueden socializar en Internet, como pedófilos, grupos sectarios, grupos de ideologías radicales, grupos violentos y mantenerlos alejados de los contenidos que producen estos grupos de personas.

Además este tipo de mecanismos pueden adaptarse para su uso en la publicidad, segmentando los anuncios en televisión o aparadores interactivos, ofreciendo cigarrillos o licor a personas adultas y el último juguete de novedad a niños.

En caso concreto permitirá al laboratorio de análisis forense digital de la Facultad de Ciencias de la Computación de la Benemérita Universidad Autónoma de Puebla implementar un detector de sitios web que almacenen y distribuyan contenido relacionado con pedofilia.

## **Hipótesis**

### **Hipótesis General**

Estimar la edad de una persona en una imagen digital usando como referencia exclusivamente el rostro.

### **Hipótesis Específicas**

- El rostro de una persona puede ser identificado en una imagen.
- Se pueden minimizar las características en las imágenes que no aportan información importante al clasificador.
- Se puede crear un clasificador basado en métodos estadísticos.
- Se puede diferenciar la mayoría o minoría de edad en una imagen facial.

## **Objetivos**

### **Objetivo General**

Desarrollar un sistema clasificador de edad que utilice una imagen digital usando como referencia exclusivamente el rostro.

## **Objetivos Específicos**

- Automatizar el preprocesamiento de imágenes porque se requiere la carga de grandes grupos de imágenes.
- Detectar el rostro en las imágenes para procesar la imagen en caso de que sí exista o descartarla en caso de que no.
- Automatizar el procesamiento de las imágenes para eliminar características que no aportan información a los modelos de clasificación.
- Creación de los modelos para entrenar el clasificador.
- Verificar resultados.

## **Resumen capítulo 1**

En el capítulo uno se abordó la importancia de generar un sistema clasificador de mayoría y minoría de edad debido a los riesgos que existen para los menores de edad en Internet así como las hipótesis generadas y los objetivos que se persiguen para resolver la problemática presentada.

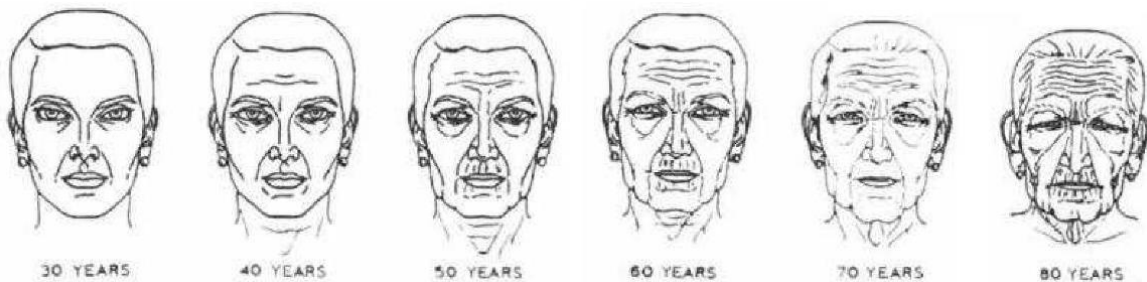
## Capítulo 2

### Marco conceptual

Esta sección aborda los diferentes rasgos característicos del envejecimiento humano y los principales problemas que existen al estimar la edad de una persona a partir de su rostro.

### Envejecimiento

En la figura 1 se muestran los principales signos de envejecimiento de un rostro como son el crecimiento craneofacial, envejecimiento de la piel debido a la aparición de arrugas, manchas y reducción de la elasticidad del músculo, además el rostro sufre ligeras variaciones apreciables en el contorno de ojos, nariz y boca [14].



**Figura 1. Esquemas de la edad de un rostro desde los 30 a los 80 años (Fu. Guo & Huang, 2010).**

Al avanzar la edad, la apariencia del rostro humano desarrolla considerables cambios, sin embargo los signos del envejecimiento son incontrolables y personalizados, sumando a ello, existen otros factores que influyen en el envejecimiento como lo son el estilo de vida (consumo de cigarrillos, alcohol, estrés), la genética o los efectos de la radiación ultravioleta en la piel. Además, factores externos como las cirugías estéticas, figura 2 o la iluminación y el maquillaje, figura 3.



**Figura 2. Efectos de la cirugía estética (Troell, 2014).**



**Figura 3. Efecto del maquillaje (Begoña, 2013).**

Analizando estos factores, se afirma que la estimación de la edad a partir de un rostro es una tarea complicada cuando la pretende realizar una persona y obtiene mayor dificultad cuando se intenta realizar de manera automática mediante un ordenador ya que es complicado definir un patrón para representar los signos de envejecimiento en una edad en específico [15].

## **Estimación de edad**

La estimación de edad tiene como objetivo etiquetar un rostro a una edad exacta o bien dentro un grupo de edades. Debido a que la primera opción es más precisa también sufrirá de mayor número de errores. Para entender la dificultad a la hora de estimar la edad de un rostro hay que diferenciar entre los siguientes conceptos sobre la edad humana [16].

- Edad real: Edad de un individuo, años acumulados desde su nacimiento.
- Edad aparente: Edad mostrada por la apariencia visual de un rostro.
- Edad percibida: Edad de un individuo calculada por personas a partir de la apariencia visual.

- Edad estimada: Edad de un individuo reconocida por una maquina a partir de la apariencia visual.

La importancia de la estimación de edad por computadora sobresale por la dificultad que tiene el ser humano en esta tarea. Por ejemplo, en [17] se realizaron estudios a una serie de vendedores de alcohol a quienes les mostraron una serie de fotografías de personas de 16 años para determinar si tenían una edad igual o mayor a 18 años y tener derecho a adquirir bebidas alcohólicas, los resultados de este estudio fueron que el 38% de los chicos y el 56% de las chicas pasaron como adultos cuando en realidad no lo eran.

Investigaciones similares mostradas en [18] determinaron que se tiende a sobreestimar la edad de las personas jóvenes en aproximadamente un tercio de los casos, por lo que los vendedores de alcohol se ven obligados a pedir la identificación en la mayoría de los casos a sus clientes.

## **Marco teórico**

Como se menciona en [15] el proceso de envejecimiento es incontrolable y personalizado por lo tanto la estimación de edad puede ser representada como un problema de clasificación o como un problema de regresión. Actualmente los sistemas existentes están compuestos por dos módulos separados: un modelo de representación de características de envejecimiento y una técnica de estimación de edad [19].

## **Modelos de representación de características**

En la actualidad existen cinco enfoques que son utilizados para representar la edad a partir de imágenes faciales; A continuación se detallara cada uno de ellos.

## Modelos Antropométricos

Se basan en la teoría del desarrollo craneofacial utilizado por Young Kwon y Niels Lobo [20] que utilizan un modelo matemático para describir el crecimiento de la cabeza de una persona desde su infancia hasta la edad adulta. Dicho modelo indica que el perfil de la cabeza está determinado por un par  $(r, \theta)$ , donde  $r$  es la distancia del origen a un punto  $P$  del perfil, y  $\theta$  el ángulo que forma el segmento  $\overline{OP}$  con el eje vertical. Nombrando  $k$  a un parámetro que se incrementa con el tiempo, los crecimientos sucesivos del perfil,  $(r', \theta')$  se modelan como se muestra en la ecuación 1:

$$\begin{cases} \theta' = \theta \\ r' = r (1 + k (1 - \cos \theta)) \end{cases} \quad (1)$$

El modelo ha dejado de utilizarse para estimar edades por los siguientes motivos.

1. El modelo solo puede describir el perfil de la cabeza en menores de edad.
2. En ocasiones el perfil de la cabeza es difícil de extraer de imágenes de caras en 2D [21].

Sucesivos trabajos de estimación de edad han utilizado la llamada antropometría de la cara, que se basa en la extracción y estudio de tamaños y proporciones en la misma. En [22] Farkas definió la antropometría en función de medidas de 57 puntos de interés en rostros de personas de diferentes edades. Para caracterizar la edad de una persona se miden las distancias entre algunos de los puntos y se consideran las proporciones entre ellas.

La idea es utilizada por Kwon y Lobo en [20] y en [23] para distinguir bebés de adultos, extrayendo 6 proporciones diferentes como se muestra en la figura 4.

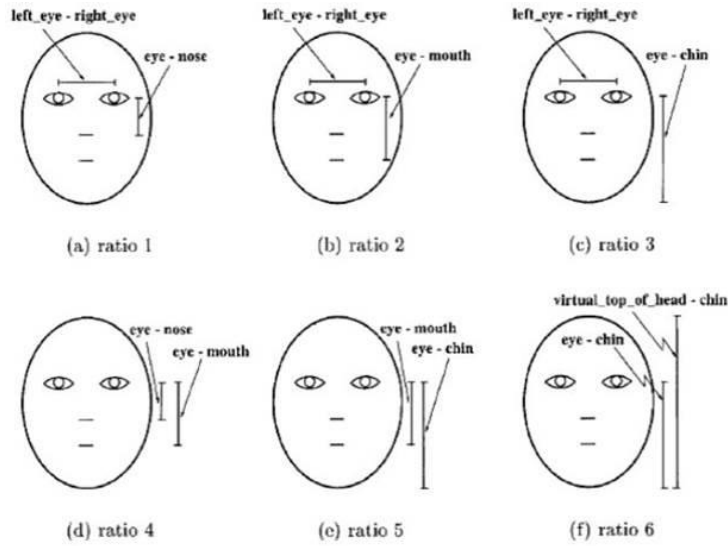


Figura 4. Modelo antropométrico (Kwon & Lobo 1994,1999).

Ramanathan y Chellappa usaron la misma idea en [21] utilizando 8 proporciones para modelar la progresión de edades comprendidas entre 0 y 18 años como se muestra en la figura 5.

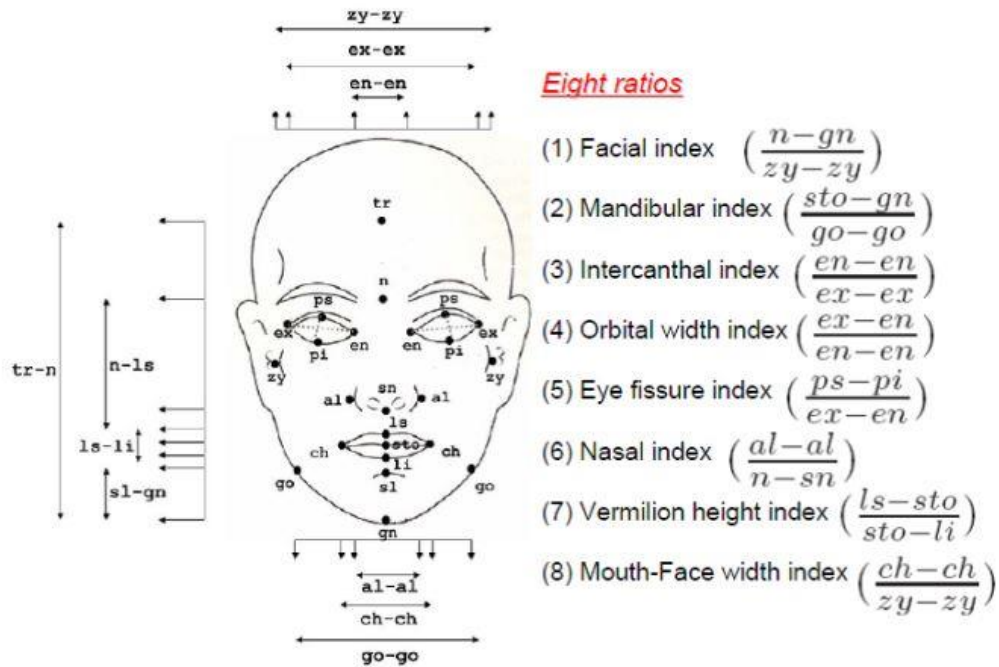


Figura 5. Modelo antropométrico (Ramanathan y Chellappa 2006).

Sin embargo, los modelos antropométricos presentan varios inconvenientes:

- Solamente son útiles en la estimación de la edad de personas jóvenes ya que los mayores cambios en la forma del cráneo se producen antes de la mayoría de edad.
- La detección de los puntos de interés en las imágenes es sensible a la pose de la cabeza, lo que dificulta su aplicación en entornos reales.
- Se requiere una adaptación a personas de otros orígenes como africanos y asiáticos debido a que las medidas dadas por Farkas solo son válidas para personas de origen caucásico (europeos o descendientes).

### **Apariencia Activa**

Los modelos de apariencia activa se basan en modelos estadísticos y fueron propuestos inicialmente para codificar imágenes de rostros [24].

Dado un conjunto de imágenes faciales y los correspondientes puntos de referencia, se genera un modelo estático de variación basado en la forma del rostro utilizado para extraer el vector de características, como se muestra en la ecuación 2

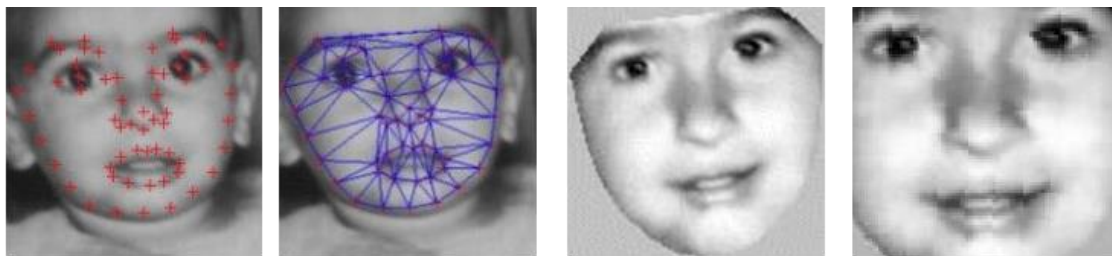
$$x_i = \bar{x} + Px_i \quad (2)$$

Donde  $x_i$  representa la forma de la imagen de entrenamiento,  $\bar{x}$  es el valor medio de las imágenes,  $P$  es la matriz de autovectores generada y  $x_i$  es el vector de características.

Los modelos de apariencia activa han sido utilizados en múltiples ocasiones para codificar rostros de manera satisfactoria. Lanitis et al en [25] extendieron los modelos de apariencia activa para describir la relación entre la edad de los individuos y la descripción de las imágenes. En

comparación con los modelos antropométricos, los modelos de apariencia activa pueden tratar con cualquier edad, en vez de tan solo con los jóvenes. Además, consideran tanto la forma como la textura en vez de las características geométricas del rostro y son robustos frente a posturas cosa que los métodos antropométricos no. Estas técnicas son aplicables a la estimación precisa de la edad, en cuanto a que cada imagen de test es etiquetada con un valor específico escogido de un rango continuo.

Las características extraídas mediante los modelos activos de apariencia, por lo general funcionan para estimar edad. Sin embargo, dichos sistemas necesitan contar con un etiquetamiento de una serie de puntos de la cara como se muestra en la figura 6, algo que debe realizarse de manera manual u obtener una base de imágenes que cuente con el etiquetado evitando que se automaticen los procesos de creación de los modelos.



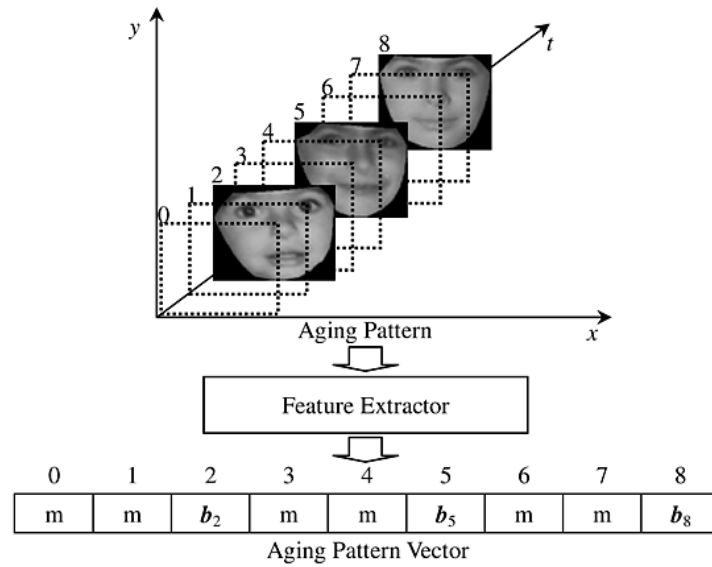
**Figura 6. Modelos de apariencia activa (Alejandro Torrisi, 2015).**

### **Patrones de envejecimiento**

Los patrones de envejecimiento son utilizados por Xin Geng, Zhi-Hua Zhou y Kate Smith-Miles en [26] donde los tratan como una secuencia de imágenes de los individuos con diferentes edades usadas de manera conjunta para modelar el proceso de envejecimiento.

Se define como patrón de envejecimiento a la secuencia de imágenes de una misma persona ordenadas temporalmente como se observa en la figura 7 y para construirla se requiere de un

subespacio que capture la variación principal en el conjunto de datos de entrenamiento utilizando el análisis de componentes principales.



**Figura 7. Patrones de envejecimiento (Xin Geng, 2007).**

Para realizar la proyección al subespacio se procede como indica la ecuación 3:

$$y = W^t (x - \mu) \quad (3)$$

Donde  $\mu$  es el valor medio de  $x$ , que es el vector que contiene las características de cada sujeto, es decir, el vector de envejecimiento y  $W^t$  es la transpuesta de  $W$  que se compone de los autovectores de la matriz de covarianza de  $x$ .

Para realizar la estimación de edad, la imagen de prueba es evaluada y asignada a un determinado patrón de envejecimiento en una posición concreta, comparando con cada una de las imágenes que forman el patrón, seleccionando la que supone un error menor en la reconstrucción.

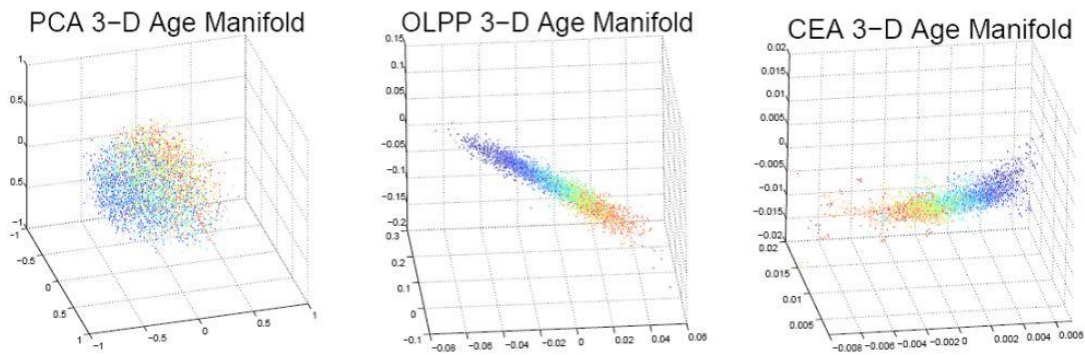
Los modelos presentan varios inconvenientes:

- Es complicado obtener el conjunto de imágenes de diferentes edades de la misma persona.
- Se supone que todas las personas tienen las mismas características faciales en determinada edad, lo cual es falso ya que el proceso de envejecimiento es individual y personalizado como indica [15].

### **Manifold de edad**

Basado en la idea de que en vez de aprender un patrón de envejecimiento para cada individuo, se puede modelar un único patrón para varios individuos a lo largo de diversas edades. Así, cada edad se representaría utilizando diferentes rostros de diferentes personas. Haciendo flexible la estimación de edad en comparación con el método del subespacio del patrón de envejecimiento. El manifold de edad [27] es una forma de modelar el patrón de edad común. Dicho método no necesita un número mínimo de imágenes de un individuo determinado, de hecho, permite que una persona solo aporte una imagen de una edad específica, lo que es práctico para crear bases de imágenes. Si visualizamos cada imagen como un punto en un espacio de  $D$  dimensiones, el objetivo del manifold de edad es encontrar una representación de datos en una dimensión  $d$  más pequeña capturando a la vez la distribución intrínseca de los datos y su estructura geométrica subyacente. En la literatura se han propuesto diversas técnicas, como OLPP (Orthogonal Locality Preserving Projections) propuesta por Deng Cai, Xiaofei He, Jiawei Han y Hong-Juang Zhang en [28], CEA (Conformal Embedding Analysis) desarrollada por Yun Fu y Thomas Huang en [29], LARR (Locally Adjusted Robust Regression) propuesta en 2008 por Goudong Guo, Yun Fu, Charles R. y Thomas Huang [30] o SSE (Synchronized Sub manifold Embedding) desarrolladas en 2009 por Shuicheng Yan, Huan Wang, Yun fu, Xiaou Tang y Thomas Huang en [31]. Los

métodos OLPP y CEA representan de mejor manera el conjunto de datos que las técnicas clásicas de reducción de dimensionalidad como se observa en la figura 8, mientras que LARR y SSE obtienen un error medio absoluto de 5.8 años en la estimación de edad.



**Figura 8. Manifold de edad (Goudong Guo, 2007).**

## **Apariencia**

Combinan la descripción de la textura y la forma del rostro y debido a que son menos sensibles a variaciones de luz, puntos de vista y rotación del rostro son más utilizados que las técnicas anteriormente descritas para representar las características de edad.

El método se basa en obtener un conjunto de rostros de entrenamiento con los cuales se define un rostro promedio de cada clase para extraer las características mostradas en la figura 9 que permiten describir la distribución de datos de cada clase.

La clasificación se obtiene al extraer las características de la imagen de prueba y compararlas con las clases creadas anteriormente.



**Figura 9. Modelos de apariencia [imagen] disponible en:**

**<http://www.adelantandoelmundo.com/2015/04/averiguar-edad-persona.html> [recuperado 08 sep 2017].**

En 2006 Ahonen, Hadid & Pietikainen en [32] desarrollaron un detector de rostros con descriptores locales (Local Binary Pattern) asignando un peso a cada parte del rostro que detectaba obteniendo un 97% de exactitud, Sanhueza Riveros en [33] y Hayashi, Yasumoto e Ito en [34] usaron este modelo para clasificar edad, buscando los elementos más significativos del rostro como son boca y ojos generando un clasificador de 4 diferentes grupos (niños, adolescentes, adultos y ancianos).

Con el paso del tiempo, se han hecho estudios con evidentes variaciones en la apariencia como Spatially Flexible Path desarrollado por Yan, Zhou, Liu, Hasegawa-Johnson y Huang en 2008 que utiliza la información de las regiones locales de las imágenes y su posición, dando más importancia a características discriminantes.

Otro método fue el desarrollado por Jinli Suo, Feng Min, Song Zhu, Shiguang Shan y Xilin Chen [35] el cual busca aplicar un filtro con diferentes parámetros a cada característica facial (ojos, nariz, boca).

Las Fisherfaces son una combinación de las técnicas estadísticas análisis de componentes principales y análisis discriminante lineal método propuesto por Belhumeur, Hespanha y Kriegman en [36], Terradez Gurrea indica en [37] que el análisis de componentes principales es una técnica estadística de síntesis de la información, es decir, ante un banco de datos con muchas variables, el objetivo es reducirlas, perdiendo la menor cantidad de información posible.

Por otro lado en análisis discriminante lineal maximiza la razón de dispersión entre clases a dispersión dentro de las clases. La idea es que las mismas clases deben de agruparse, mientras que las clases diferentes están tan lejos como es posible en una representación de baja dimensión, como lo indica Hernández y Méndez en [38].

### **Fisherfaces como representación de características**

Como indica Belhumeur en [36] están orientadas a la creación de modelos que describen las características de los conjuntos de imágenes de cada clase haciendo énfasis en sus diferencias.

Para generar los modelos se utiliza un conjunto de imágenes asignando una cantidad similar de imágenes para cada clase, después se aplica sobre ellas los métodos estadísticos: análisis de componentes principales y análisis discriminante lineal.

*Análisis de componentes principales:* Obtiene una proyección lineal y maximiza la dispersión de todas las imágenes proyectadas.

En primer lugar, se considera el conjunto de  $N$  imágenes con valores en el espacio de imágenes  $n$ -dimensional. Como lo indica la ecuación 7.

$$\{x_i\} \quad i = 1, 2, \dots, N \quad (4)$$

Se asume que cada una de las imágenes pertenece a una de las  $c$  clases  $\{x_1, x_2, \dots, x_c\}$ . Asimismo, se considera una transformación lineal que lleva el espacio de imágenes original de  $n$  dimensiones al espacio de características de dimensión  $m$ , donde  $m < n$ . Los nuevos vectores de características  $Y_k \in \mathfrak{R}_m$  son definidos por la siguiente transformación lineal, ecuación 8.

$$Y_k = W^T X_k \quad K = 1, 2, \dots, N \quad (5)$$

Donde  $W \in \mathfrak{R}^{n \times m}$  es una matriz con columnas ortonormales.

Se define la matriz de distribución total  $S_T$  como indica la ecuación 9.

$$S_T = \sum_{K=1}^N (x_k - \mu)(x_k - \mu)^T \quad (6)$$

Donde  $\mu \in \mathfrak{R}_n$  es la media de todas las imágenes de (4). Luego de aplicar la transformación lineal  $W_T$ , la distribución de los vectores de características  $\{y_1, y_2, \dots, y_N\}$  es  $W_T S_T W_T$ . Se utiliza la proyección  $W_{opt}$  que maximiza el determinante de la distribución total de la matriz de las  $N$  imágenes. Ecuación 7.

$$W_{opt} = \arg \max W |W^T S_T W| = [w_1, w_2 \dots w_m] \quad (7)$$

La figura 10 muestra la salida del análisis de componentes principales sobre dos dimensiones de un conjunto de imágenes que contiene gente adulta (puntos rojos) y gente joven (puntos amarillos) y donde se observa que la representación de datos de cada imagen ha sido reducida del total de píxeles a dos vectores que componen la dirección donde la imagen tiene más diferencia.

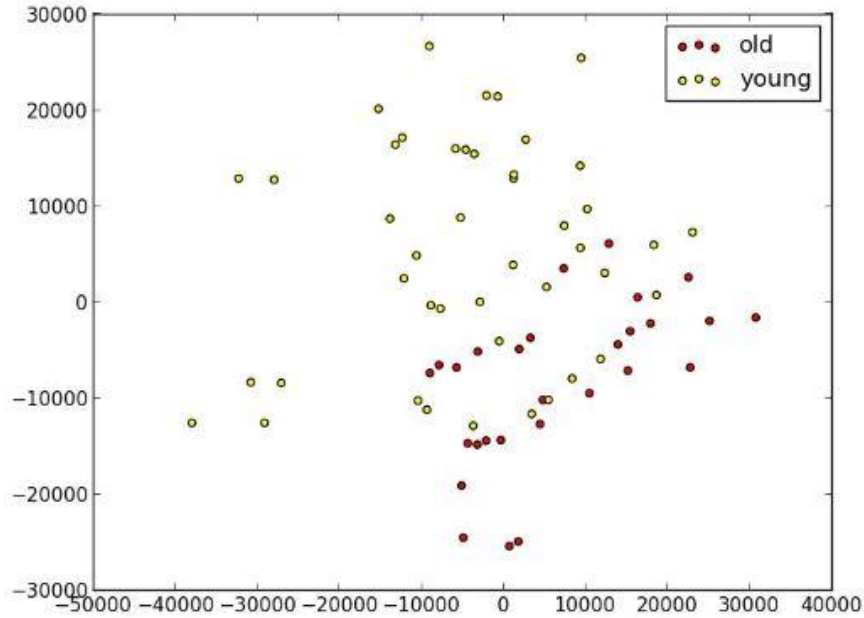


Figura 10. Gráfico de valores vectoriales para imágenes de ancianos y jóvenes (Patterson 2014).

**Análisis discriminante lineal:** Después de que los datos han sido proyectados por el análisis de componentes principales, se rotan los ejes en buscando minimizar la distancia entre miembros de la misma clase, mientras maximiza la distancia entre clases.

Para ello este método selecciona el  $W$  de la ecuación 7, de manera que el cociente entre la distribución de las diferentes clases y la distribución intraclases sea lo más alta posible.

Para esto, se define la matriz  $S_B$  de distribución entre clases como lo indica la ecuación 8.

$$S_T = \sum_{i=1}^c N_i (\mu_i - \mu) (\mu_i - \mu)^T \quad (8)$$

Y la matriz  $S_W$  de distribución intra-clases, como se puede apreciar en la ecuación 9.

$$S_W = \sum_{i=1}^c \sum_{x_k \in X_i} N_i (\mu_i - \mu) (\mu_i - \mu)^T \quad (9)$$

Donde  $\mu_i$  es la imagen media de la clase  $X_i$ , y  $N_i$  es el número de imágenes en la clase  $X_i$ . Si la matriz  $S_W$  es no singular, la proyección  $W_{opt}$  se elige como la matriz con columnas ortonormales que aumenta el cociente del determinante de la matriz de distribución entre clases de las imágenes proyectadas y el determinante de la matriz de la distribución intraclases de las  $N$  imágenes, de la ecuación 10.

$$W_{opt} = \arg \max W \frac{|W^T S_B W|}{|W^T S_W W|} = [w_1, w_2 \dots w_m] \quad (10)$$

Donde  $\{W_i \mid i = 1, 2, \dots, m\}$  es el conjunto de valores propios de  $S_B$  y  $S_W$  correspondiente a los  $m$  mayores valores propios  $\{\lambda_i \mid i = 1, 2, \dots, m\}$ , entonces, se muestra de la ecuación 11.

$$S_B W_i = \lambda_i S_W W_i \quad i = 1, 2, \dots, m \quad (11)$$

Por lo tanto, se obtienen  $c - 1$  valores propios distintos de cero, y por consiguiente el límite superior de  $m$  es  $c - 1$ , donde  $c$  es el número de clases, ecuación 12.

$$W_{pca}^T = W_{lda}^T W_{pca}^T \quad (12)$$

Donde se derivan las ecuaciones 13 y 14.

$$W_{pca} = \arg \max W |W^T S_T W| \quad (13)$$

$$W_{lda} = \arg \max W \frac{|W^T W_{pca}^T S_B W_{pca} W|}{|W^T W_{pca}^T S_W W_{pca} W|} \quad (14)$$

En la figura 11 se muestra una gráfica ideal creada por el análisis discriminante lineal.

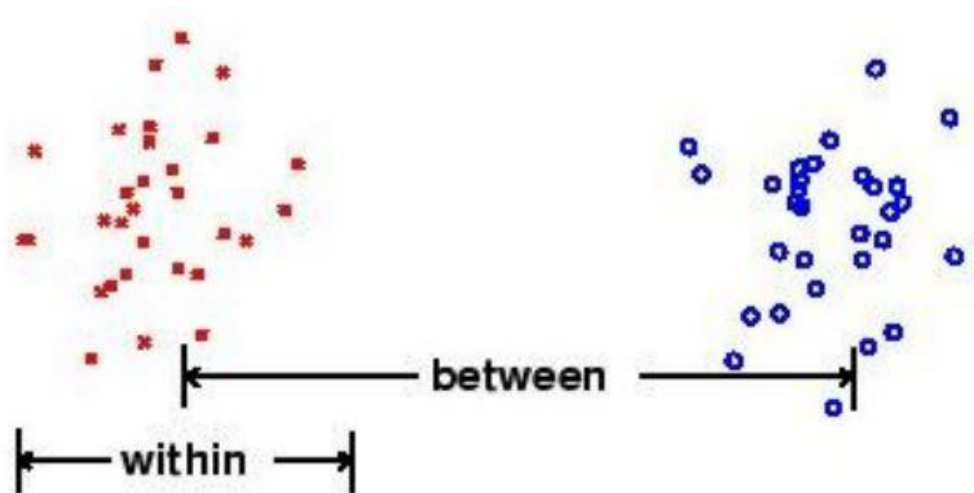


Figura 11. Ejemplo de una gráfica ideal creada por el análisis discriminante lineal [imagen] disponible en: <http://www.dtrek.com/LdaSepGood.jpg> [recuperado 2017-09-20].

En la figura 12 se observa un mapa de colores de las características que separan ambas clases donde los píxeles rojos son los rasgos más distintivos.

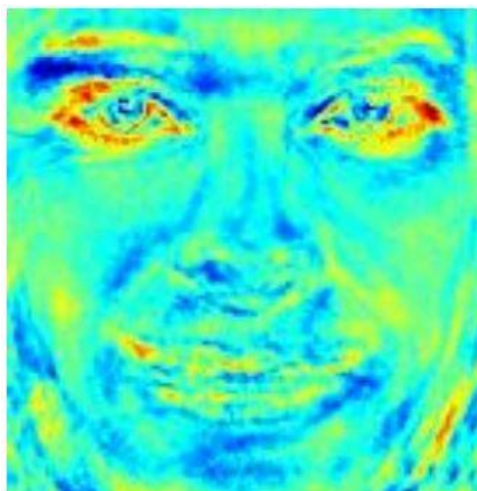


Figura 12. Ejemplo de salida de Fisherface detallando las características de la clase adultos [imagen] disponible en: [http://docs.opencv.org/trunk/\\_images/fisherface\\_0.png](http://docs.opencv.org/trunk/_images/fisherface_0.png) [recuperado 2017-09-20].

De esta manera el modelo de apariencia permite obtener las características de imágenes faciales generando los clasificadores con un conjunto de imágenes de entrenamiento con una descripción previa. En el artículo documental [40], Kyle Patterson presenta una propuesta para crear

exhibidores interactivos para museos, los cuales muestran información segmentada para diferentes tipos de visitantes, en él se utilizan las Fisherfaces como método para obtener las características de las imágenes faciales haciendo énfasis en la separación entre clases.

## **Técnicas de estimación de edad**

La estimación de edad es vista como un proceso de reconocimiento de patrones, donde cada uno de los descriptores de la edad es representada por una clase y puede abordarse con técnicas como la clasificación por vecindad, redes neuronales o máquinas de soporte vectorial, por otro lado las edades son un conjunto de números con valores secuenciales, por lo tanto la estimación de edad también puede ser vista como un problema de regresión [41].

## **Clasificación por vecindad**

Es un algoritmo de aprendizaje supervisado en el que dado unos vectores de entrenamiento en un espacio multidimensional, para un nuevo vector se calcula la distancia entre los vectores de entrenamiento y aquél y se seleccionan los  $k$  más cercanos [42]. La medida de distancia más utilizada en este método es la distancia euclídea:

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^p (x_{i_r} - x_{j_r})^2} \quad (15)$$

Donde los valores de los atributos del  $i$ -ésimo ejemplo donde ( $1 \leq i \leq n$ ) se representan por el vector  $p$ -dimensional  $x_i = (x_{i_1}, x_{i_2}, \dots, x_{i_p}) \in X$ .

En [43] Khaung Tin, utilizo este algoritmo para clasificar la edad utilizando 7 grupos específicos reduciendo el número de características con pca y obteniendo un 72.58% de exactitud.

### Máquina de soporte vectorial

Es un sistema de aprendizaje automático que permite resolver problemas de clasificación y regresión de manera eficiente ya que están poseen una sólida fundamentación matemática y se basan en el concepto de minimización del riesgo estructural.

Una máquina de soporte vectorial mapea los puntos de entrada a un espacio de características de mayor dimensión, por ejemplo los puntos de entrada están en  $R^2$  entonces son mapeados a  $R^3$  encontrando un hiperplano que separe y maximice el margen  $m$  entre las clases como se observa en la figura número 13.

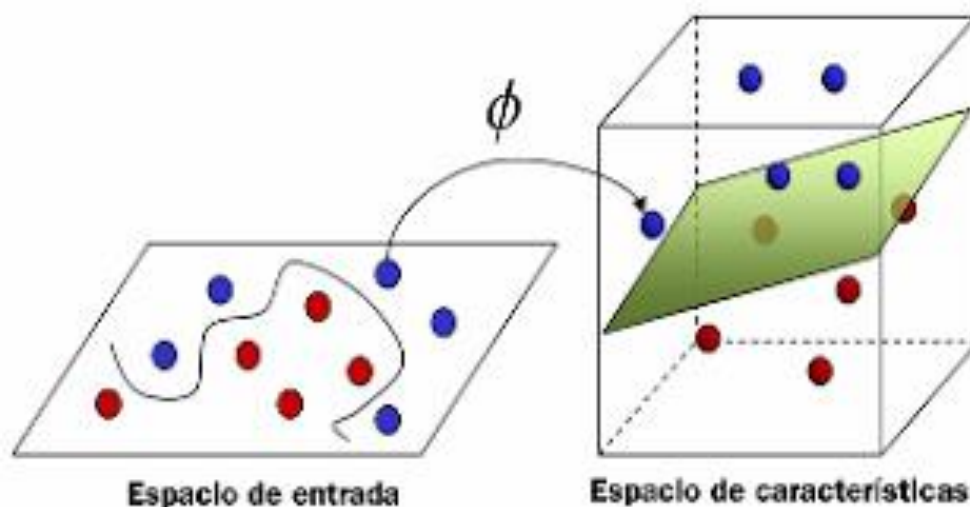


Figura 13. Máquinas de soporte vectorial (Canales, 2009).

En [44] Gustavo Betancourt indica que dado un conjunto  $S$  de puntos etiquetados como se observa en la ecuación 16:

$$(y_1, x_1) \dots (y_i, x_i) \tag{16}$$

Cada punto de entrenamiento  $x_i \in R^N$  pertenece a alguna de dos clases y se le ha dado una etiqueta  $y_i \in \{-1,1\}$  para  $i = 1, \dots, 1$ . En la mayoría de los casos, la búsqueda de un hiperplano adecuado en un espacio de entrada es restrictivo para ser de uso práctico.

Una solución a esta situación es mapear el espacio de entrada en un espacio de características de una dimensión mayor y buscar el hiperplano óptimo.

Sea  $Z = \varphi(x)$  la notación del correspondiente vector en el espacio de características con un mapeo  $\varphi$  de  $R^N$  a un espacio de características  $Z$ , como se observa en la ecuación 17:

$$w * z + b = 0 \tag{17}$$

Definido por el par  $(w, b)$ , tal que podamos separar el punto  $x_i$  de acuerdo a la ecuación 18:

$$f(x_i) = \text{Sign}(w * z + b) = \begin{cases} 1 & y_i = 1 \\ -1 & y_i = -1 \end{cases} \tag{18}$$

Donde el conjunto  $S$  es linealmente separable si existe  $(w, b)$  tal que las inecuaciones 19

$$\begin{cases} (w * z_i + b) \geq 1, & y_i = 1 \\ (w * z_i + b) \leq -1, & y_i = -1 \end{cases} \tag{19}$$

Sean válidas para todos los elementos del conjunto  $S$ , se puede encontrar un único hiperplano óptimo, para el cual, el margen entre las proyecciones de los puntos del conjunto  $S$  de dos diferentes clases es maximizado.

En [45] se desarrolló un método para estimar la edad a partir del rostro llamado BIF (Biologically Inspired Features), la idea es que el modelo contiene capas alternadas, unas simples (S) y otras

complejas (C). La complejidad de las capas va aumentando según nos movemos desde el córtex visual primario al córtex inferior temporal. La primera capa, S1, se crea con filtros Gabor sobre la imagen de entrada. La segunda capa C1, se obtiene aplicando una operación MAX sobre la capa S1. Este tipo de características tienen la ventaja de que son robustas a pequeñas translaciones, rotaciones, escalado, capturando con eficacia los patrones de edad. En el mismo artículo utilizaron las máquinas de soporte vectorial, donde se obtuvo un error absoluto medio de 4.77 años para la base de imágenes FG-Net y de 3.82 con la base de imágenes YGA.

### **Redes neuronales**

Fueron diseñadas para emular las características del funcionamiento básico de las neuronas biológicas. Consiste en aplicar un conjunto de entradas a una neurona y mediante una combinación línea se representa una salida conectada a la entrada de otra neurona. Cada entrada está asociada a un peso individual o ponderación correspondiente al análogo del grado de conexión de la sinapsis. Luego, el resultado de la combinación entregará un valor, denominado grado de certeza entre 1 y -1 permitiendo a las neuronas calcular funciones lógicas como AND, OR o NOT. Una red neuronal (conjunto de neuronas) puede ajustar los pesos para asociar a cada entrada el valor deseado de salida. Esta habilidad de adaptación se utiliza en problemas donde hay un número finito de salidas, pero no una manera realista de representar todas las posibles entradas. En la figura 14 se puede ver una representación gráfica de una neurona artificial.

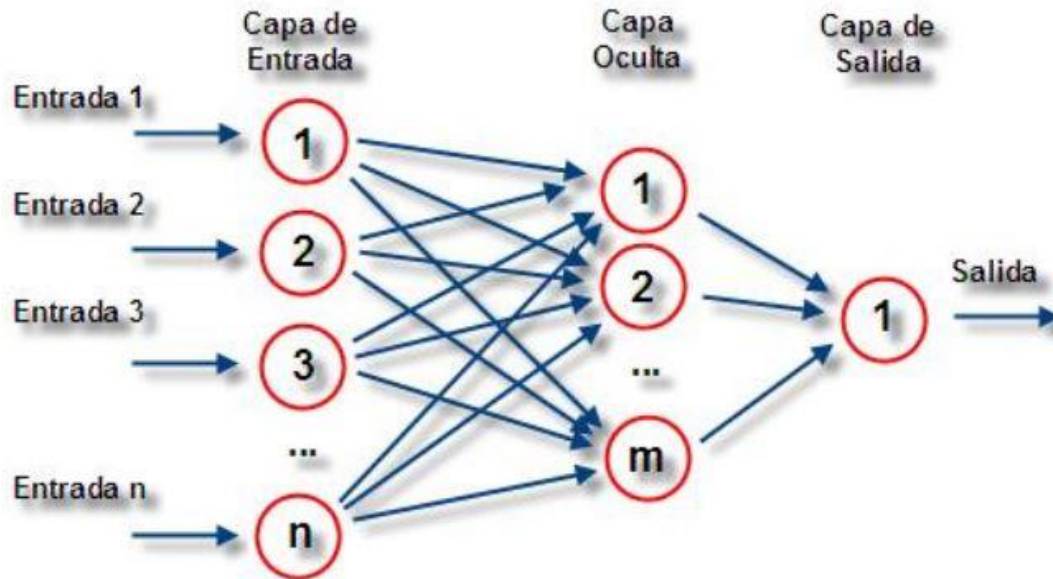
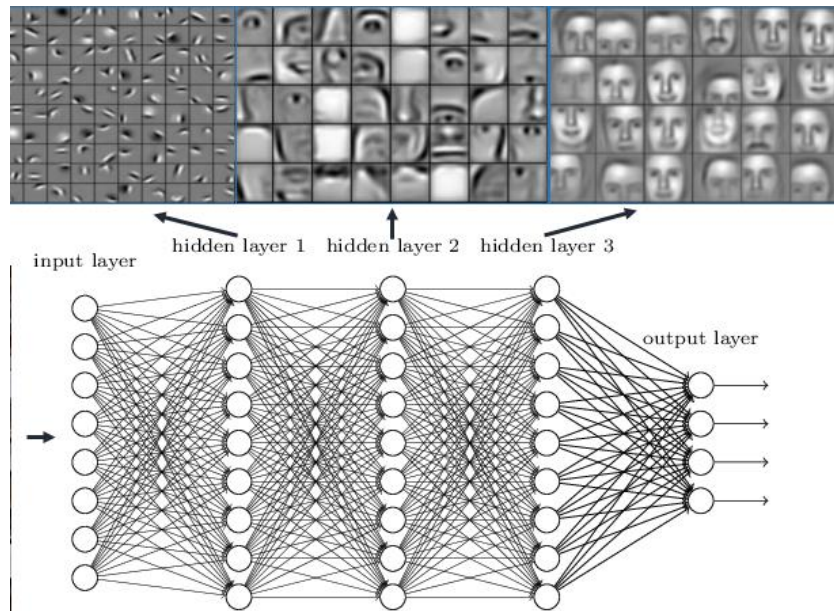


Figura 14. Redes neuronales (González-Rodríguez, A. F. 2014).

La capacidad de cálculo y potencia de la computación neuronal proviene de las múltiples conexiones de las neuronas artificiales. La red más simple es un grupo de neuronas ordenadas en una capa. Los nodos circulares son distribuidores de las entradas y no se consideran constituyentes de una capa.

Las redes multicapa se forman con un grupo de capas simples en cascada. La salida de una capa es la entrada de la siguiente capa. Se ha demostrado que las redes multicapa presentan cualidades y aspectos por encima de las redes de una capa simple.

En el contexto de la clasificación de edad, se crea una red neuronal en una estructura que contiene 4 capas como se indica en la figura 15. La primera capa consiste en neuronas que son responsables de ingresar la información del rostro a la red. La segunda y tercera capas se encuentran ocultas y permiten a la red reducir los para obtener la salida esperada. La última capa es en donde cada salida posible es representada por una neurona en particular [46].



**Figura 15. Redes neuronales. (2017). [imagen] disponible en: <http://www.rsipvision.com/exploring-deep-learning/> [recuperado 8 Sep. 2017].**

Las redes neuronales son utilizadas por Jinli Suo et al. En [35] con un perceptron multicapa y por Hewahi en [47], para clasificar en 4 grupos principales (niño, joven, adulto, viejo) usando como conjuntos de entrenamiento las bases de datos de FG-NET y empleando 4 capas de redes neuronales lograron obtener hasta un 78% de exactitud total con un porcentaje 87.6%, 67%, 86.4% y 70.9% respectivo a cada grupo principal.

### **Modelos de regresión**

En un escenario de estimación de edad planteado como un problema de regresión Andreas Lanitis, Chris Taylor y Timothy Cootes proponen en [48] una función de edad con tres formulaciones diferentes (lineal, cuadrática y cubica) donde los parámetros del modelo se ajustan con un algoritmo genético.

Basados en este modelo Fu y Huang proponen en [49] un modelo de regresión múltiple obteniendo significativas mejoras en rendimiento respecto a otras técnicas.

## **Regresión de soporte vectorial**

La técnica de los vectores soporte es una herramienta universal para resolver problemas de estimación de funciones multidimensionales. La idea es seleccionar el hiperplano regresor que mejor se ajuste a nuestro conjunto de datos de entrenamiento. La idea se basa en considerar una distancia margen  $\epsilon$ , de modo que todos los ejemplos se encuentren en una banda o tubo entorno a nuestro hiperplano, es decir, que disten una cantidad menor de  $\epsilon$  del hiperplano. Cuando se define el hiperplano sólo se consideran los ejemplos que disten más de  $\epsilon$  de nuestro hiperplano. En este caso esos ejemplos serán los considerados como vectores soporte [50], esta técnica se ha empleado con éxito en los problemas de estimación de edad en artículos como [51] y en sistemas como [52] donde después de tratar las características de las imágenes con diferentes algoritmos, aplicaban regresión mediante SVR mejorando los resultados con anteriores sistemas.

## **Biblioteca de procesamiento OpenCV**

Apoyarse en biblioteca de procesamiento de imágenes resulta imprescindible para evitar trabajo laborioso e innecesario. Intel OpenCV9 (Intel Open Source Computer Vision Library) [53] es una biblioteca de código abierto, gratuita, multiplataforma, rápida, fácil de utilizar, en continuo desarrollo y cada vez más extendida entre la comunidad de visión artificial. En el dominio específico de los rostros humanos, contiene algunas funcionalidades avanzadas para la detección y seguimiento del mismo. Además, se solucionan algunos aspectos auxiliares como la lectura, captura de la cámara y escritura de imágenes y vídeo.

## **Resumen capítulo 2**

En la primera parte del segundo capítulo se abordó la terminología de envejecimiento y estimación de edad permitiendo conocer las características más representativas así como considerar los principales problemas que tienen los seres humanos y los computadores al intentar estimar la edad.

En la segunda parte se explicó que un sistema clasificador de edad está construido por dos partes; los modelos de representación de características donde encontramos técnicas como los modelos antropométricos, la apariencia activa, los patrones de envejecimiento, el manifold de edad y por último los modelos basados en apariencia y una técnica de estimación de edad donde las más conocidas es la clasificación por vecindad, la máquina de soporte vectorial, las redes neuronales, los modelos de regresión y la regresión de soporte vectorial.

Además se explicó como las Fisherfaces que son la combinación de dos métodos estadísticos sirven como modelos de representación de características y por su funcionamiento están clasificadas como modelos basados en apariencia y se expuso el algoritmo para contruirlas.

# Capítulo 3

## Metodología

El modelo general del clasificador se compone de cuatro modelos independientes adaptados como se describe en la figura 16.

1.- Imagen de entrada; Busca o se genera una base de imágenes faciales que contengan la información de la edad de las personas que aparecen en cada imagen.

2.- Preprocesamiento de las imágenes; Permite cargar un conjunto de imágenes y mediante un proceso que incluye aplicación de filtros, como resultado se genera una lista de imágenes faciales depurada y estandarizada.

3- Modelo de representación de características; Utiliza las imágenes preprocesadas y genera los modelos de representación estadística de las imágenes haciendo énfasis en las diferencias entre clases.

4.- Clasificación de la nueva imagen: Una vez que se crearon y cargaron los modelos de representación, se comparan con una nueva imagen de prueba la cual es preprocesada con los mismos filtros y se extraen sus características usando el mismo modelo de representación que las imágenes de entrenamiento con el fin de obtener la estimación del grupo etario al que pertenece el nuevo rostro.

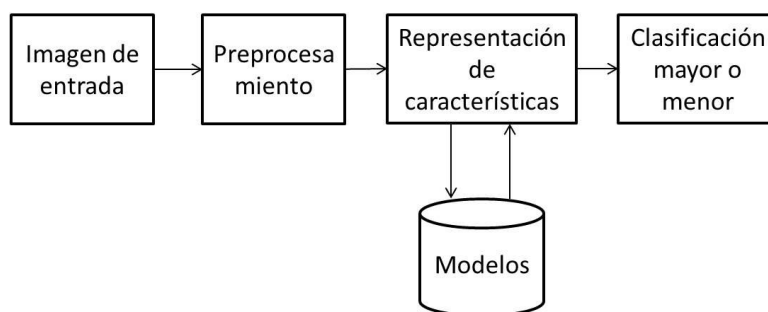


Figura 16. Modelo general del clasificador.

## **Bases de datos de envejecimiento**

Existen algunos problemas comunes en la detección de rostros que se dan en el proceso de adquisición de las imágenes, condiciones de luminosidad, bajo contraste, objetos extraños, rostros parcialmente capturados, distorsión de la imagen, etc.

Patterson en [40] indica los problemas más comunes:

- Posición; Las imágenes varían de posición y orientación y esto se debe al medio de captura o a la posición de las personas en el momento de la captura de la imagen.
- Expresión facial; Por ejemplo risa, enojo, tristeza, entre otros.
- Presencia de elementos; Barba, bigote, lentes que presentan un problema al intentar detectar un rostro.
- Oclusión; Existen objetos en la imagen que ocultan parte del rostro.

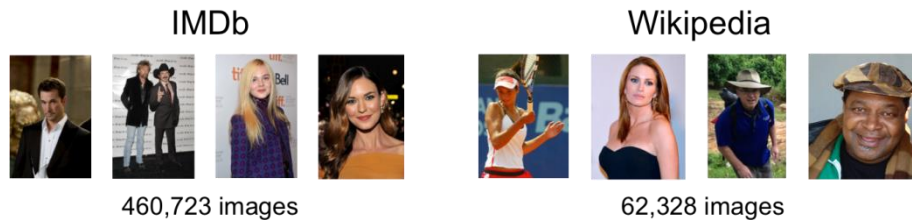
El primer paso para desarrollar un clasificador es encontrar un conjunto de imágenes de rostros que estén etiquetadas con la edad de la persona que aparece en ella.

### **IMDB-Wiki data set**

Rothe en [54] explica que IMBD-Wiki data set es una base de datos que extrae fotos de páginas web como IMDB y Wikipedia, obteniendo un total de 524,230 imágenes de rostros y con un rango de 0 a 99 años con variación de expresión, pose y luminosidad.

Dado que la base de datos IMDB-wiki, data set esta etiquetada con la fecha de nacimiento de la persona que aparece en la fotografía y la fecha en que fue tomada la foto, permitiendo obtener la edad de la personas.

En la figura 17 se observa una muestra de las base de imágenes IMBD-Wiki.



**Figura 17. Base de imágenes IMBD-Wiki.**

### **FG-Net**

FG-net base de datos de envejecimiento disponible en [55], es una base de imágenes pública que se puede utilizar para los experimentos. La base de datos contiene 1002 rostros a color y escala de grises de alta resolución con alta variación de luminosidad, pose expresión y tamaño. Contiene en total 82 sujetos de múltiples razas con un rango de edad de 0 a 69 años explicado por Guo, Fu y Dyer en [56] y es ampliamente usada como en [57] por Guo, Mu, Fu y Huang y por Luu, Ricanek, Bui y Suen en [58]. En la figura 18 se observa un ejemplo de la base de imágenes FG-Net.



**Figura 18. Ejemplo base de datos Fg-Net.**

### **Detección del rostro**

Para poder realizar este paso se utiliza el método propuesto por Viola y Jones [59] que está incluido en la Librería de código abierto Intel OpenCV [60].

El algoritmo AdaBoost busca una combinación óptima de un conjunto de clasificadores débiles con el fin de crear uno más potente. El entrenamiento consiste en un proceso iterativo, en el que se van seleccionando y dando pesos a los clasificadores elementales, según los propios ejemplos de entrenamiento.

Sea  $X_i$  las muestras de entrenamiento con  $X$  como el conjunto de todos los ejemplos y  $h_j : X \rightarrow \{+1, -1\}$  los clasificadores débiles posibles con  $J = 1, \dots, J$ ; y  $\alpha_t$  la relevancia de cada clasificador débil seleccionado con  $t = 1, \dots, T$ ; y  $T$  el número de clasificadores débiles que se quieren en el resultado. Se define una función,  $D_t(i)$  para controlar la dificultad implícita de cada ejemplo  $X_i$  en el paso  $t$ . A grandes rasgos, el funcionamiento del algoritmo AdaBoost es el siguiente:

1. Iniciar los valores  $D_1(i)$  con valores constantes, para todo  $i$
2. Para  $t = 1, \dots, T$  hacer:
  - a) Normalizar los valores de  $D_t(i)$  para que la suma total sea 1
  - b) Entrenar los clasificadores débiles  $h_j$ , para los ejemplos de  $X$ , según los pesos de  $D_t(i)$ . Para cada clasificador, calcular el error de  $h_j$  sobre los datos de prueba,  $\varepsilon_t$  (como la clasificación es binaria, el error siempre será menor que 0.5).
  - c) Seleccionar el clasificador débil,  $h_j$  que produzca el menor error  $\varepsilon_t$ .
  - d) El peso del clasificador débil  $h_t$  viene dado por  $\alpha_t = \frac{1}{2} \ln \frac{(1-\varepsilon_t)}{\varepsilon_t}$
  - e) Actualizar la función de relevancia de los ejemplos  $D_{t+1}(i)$ , disminuyendo la relevancia de los que han sido bien clasificados y aumentando los erróneos.

3. El clasificador combinado resultante es:  $H(x) = \text{signo}(\sum_t \alpha_t h_t(x))$ .

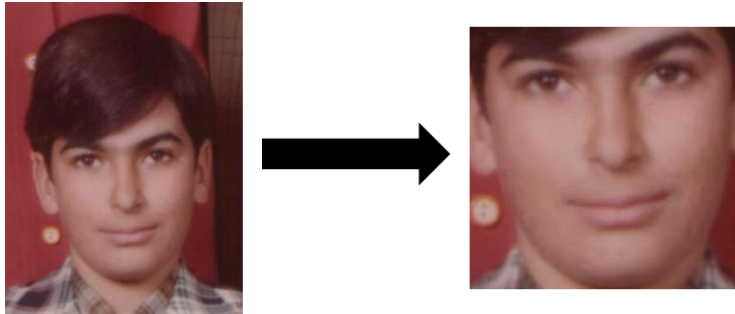
La propuesta de Viola y Jones [59], consiste de un operador definido en una subregión rectangular de la ventana, donde una parte de los píxeles se suman y los otros se restan. Si el valor del filtro está por encima de cierto umbral, se acepta, caso contrario se rechaza.

Cabe resaltar que Viola y Jones utilizan los clasificadores en un proceso en cascada, es decir se aplican de forma secuencial. Si alguno de ellos produce una respuesta negativa, se detiene el proceso y cuando toda la cascada de clasificadores devuelve un resultado positivo se indica la detección de un rostro.

En sus experimentos, los autores informan de ratios de detección ligeramente inferiores a los de otros métodos como el detector de SNoW de Yang, Roth y Ahuja [61] y superior a las redes neuronales de Rowley [62]. Sin embargo es 15 veces más rápido que el primero y 600 veces más rápido que el segundo por lo cual es ampliamente usado dentro del campo de detección de rostro y también será utilizado en esta investigación.

Con este método se puede obtener el área de interés de las imágenes y eliminar aquellas imágenes que en las que no se puede detectar el evitando insertar características innecesarias a los modelos de representación.

Una vez obtenido el área de interés se extrae y se redimensiona la imagen a un tamaño de 400 x 400 píxeles.



**Figura 19. Detección y recorte del rostro.**

## **Ruido y tipos de ruidos**

Se define como ruido a aquellos píxeles erróneos que se entremezclan de forma aleatoria con los píxeles aceptables de la imagen degradando la calidad de la imagen, El ruido puede ser añadido mediante el proceso de captura de la imagen o a través de su transmisión.

Los diferentes tipos de ruidos son:

**Ruido Gaussiano:** Es producto de pequeñas variaciones que se dan en la imagen, este tipo de ruido afecta a todos los píxeles de la imagen debido a que es producido por los componentes electrónicos como pueden ser sensores, digitalizaciones, entre otros.



**Figura 20. Imagen con ruido gaussiano.**

**Ruido sal y pimienta:** Se presenta como píxeles aleatorios que obtienen valores altos o bajos y se pueden apreciar como puntos blancos o negros. Por tal motivo se le ha dado el nombre de ruido sal y pimienta, en donde los valores máximos representan la sal y los valores mínimos la pimienta [63].



**Figura 21. Imagen con ruido sal y pimienta.**

## **Filtrado de las imágenes**

El objetivo de aplicar un filtro es obtener a partir de una imagen original, otra imagen cuyo resultado sea más adecuado para una aplicación específica, mejorando las características de la misma que permitiendo efectuar operaciones de procesado sobre ella.

Los beneficios que se obtienen al aplicar un filtro son:

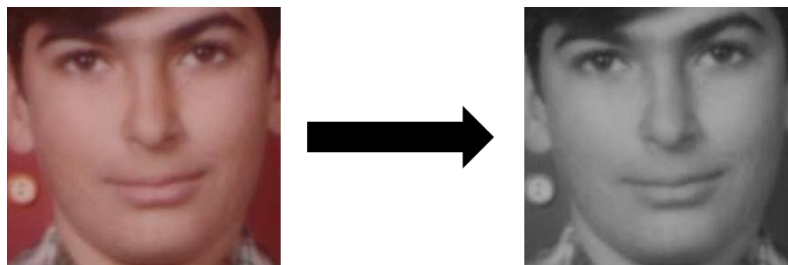
- Suavizar la imagen; Reducir la cantidad de variaciones de intensidad entre píxeles vecinos.
- Eliminar ruido; Eliminar aquellos píxeles cuyos niveles de intensidad es completamente diferente al de sus vecinos.
- Realzar bordes; Destacar los bordes que se localizan en la imagen.

En el apéndice A se muestra el resultado del clasificador utilizando distintos filtros y concluyendo que la mejor configuración es cuando se utiliza el filtro escala de grises y el ecualizado del histograma los cuales se explicaran a continuación.

El filtro escala de grises o escala monocromática utiliza la luminancia del color en los 3 componentes del color de una imagen, la aceptación general para el cálculo de la luminancia de un valor RGB es la mostrada en la ecuación 21.

$$f(x, y) = 0.299 * R + 0.587 * G + 0,114 * B \quad (20)$$

Donde R, G y B son los valores de los píxeles en cada uno de los planos de color rojo, verde y azul respectivamente, y  $f(x, y)$  es su respectivo valor a niveles de gris, esto permite reducir el costo computacional ya que evita trabajar con tres matrices una por cada color RGB y se trabaja únicamente con la matriz monocromática [64].

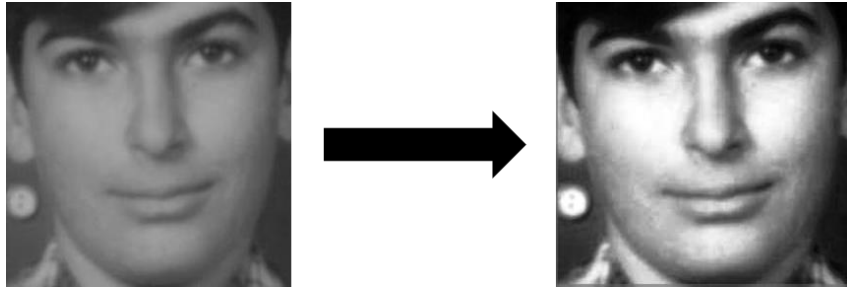


**Figura 22. Conversión de una imagen RGB a escala de grises.**

Paso siguiente paso se ecualiza el histograma obteniendo una distribución uniforme de los niveles de gris en la imagen, dado una imagen  $M \times N$  con  $n_k$  píxeles para cada nivel  $r_k$  la ecualización del histograma consiste en realizar la transformación mostrada en la ecuación 22 sobre los niveles de la intensidad de la imagen:

$$S_k = (L - 1) \sum_{j=0}^k P_r(r_j) = L - \frac{1}{MN} \sum_{j=0}^k n_j \quad (21)$$

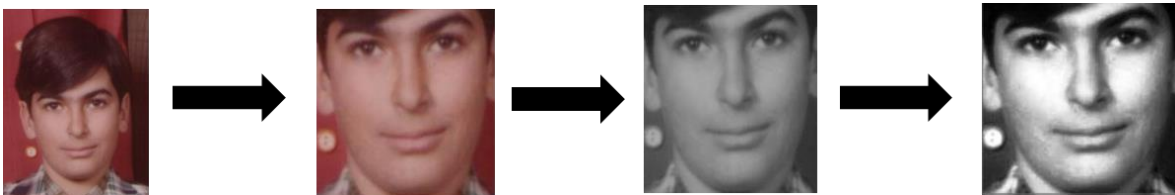
Lo que resulta en una dispersión del histograma en un rango mayor dentro del intervalo  $[0, L - 1]$ .



**Figura 23. Ecuilizado del histograma de una imagen.**

Y por último se redimensiona la imagen a un tamaño de 200x200.

Como resultado obtenemos una lista de rostros ecualizados y estandarizados. En la figura 24, se puede observar el preprocesamiento total aplicado paso a paso a un elemento de la base de imágenes FG-net.



**Figura 24. Preprocesado de una imagen.**

## Propuesta del clasificador

Se propone el desarrollo de un clasificador basado en apariencia, de acuerdo en lo documentado por Suárez Ricardo en [65], Kyle Patterson en [40] y Sanhueza Riveros en [33], debido a que este tipo de clasificadores no son sensibles a variación de pose y luminosidad, siendo este un problema recurrente cuando se emplean puntos fiduciales como lo demuestran Cox et al. En [66] o clasificadores basados en modelos antropométricos trabajados por Kwon y Lobo en [23] donde se tiene que preseleccionar las imágenes ocupando solamente las que contienen rostros completamente frontales. En los trabajos de investigación el clasificador funciona mediante dos partes específicas, la creación de los modelos de representación de características mediante Fisherfaces y la clasificación de la nueva imagen utilizando clasificación por vecindad como se puede observar en la imagen 25.

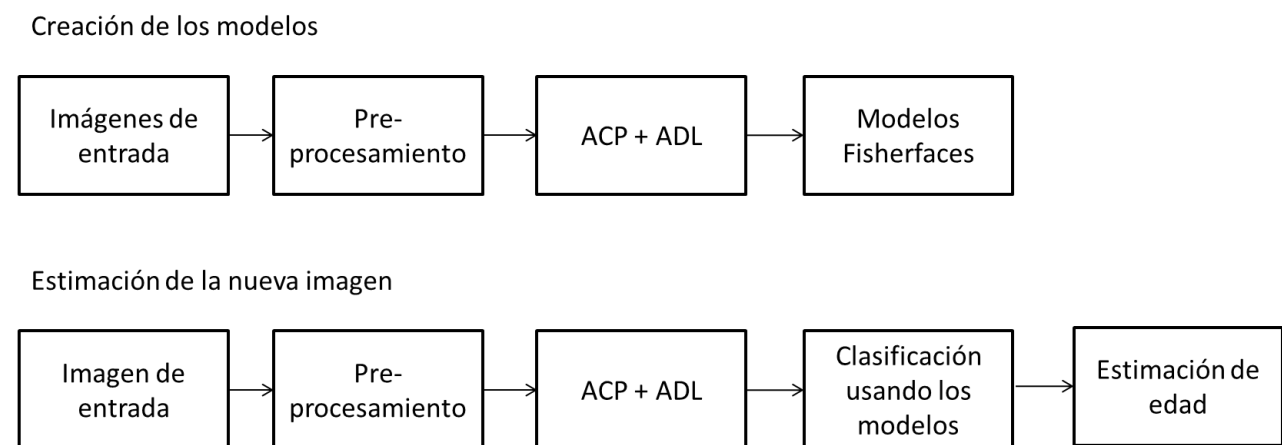
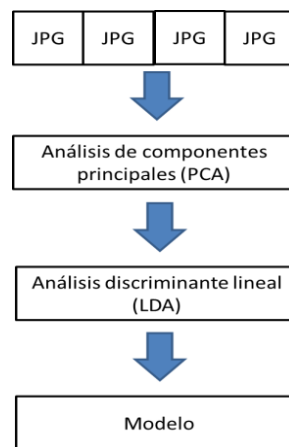


Figura 25. Módulos del clasificador.

## Creación de los modelos

Como se indica la figura 25, el clasificador está conformado por dos módulos separados, el primero está orientado a la creación de los modelos que describan las características de los conjuntos de imágenes de cada clase haciendo énfasis en sus diferencias, para generar los modelos se utiliza el conjunto de imágenes preprocesadas resultado del procedimiento anterior y se asigna una cantidad similar de imágenes para cada clase después se aplica sobre ellas los métodos estadísticos: Análisis de Componentes Principales y Análisis Discriminante Lineal como se observa en la figura 26.



**Figura 26. Creación de modelos Fisherfaces.**

El resultado de aplicar la técnica Fisherfaces sobre un conjunto de imágenes de mayores y menores de edad es el observado en la figura 27.



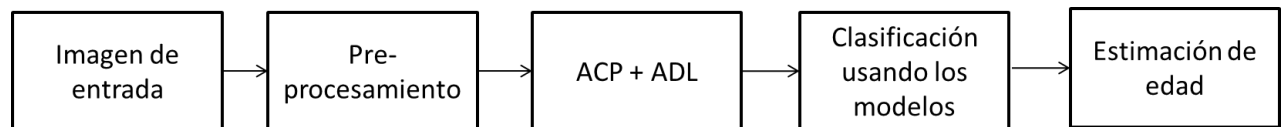
**Figura 27. Fisherface mayores y menores de edad.**

### **Clasificación de una nueva imagen**

La segunda parte de la propuesta corresponde a la clasificación de las nuevas imágenes, para poder clasificarlas se debe usar el mismo preprocesamiento:

1. Obtener el área de interés (detectar y recortar el rostro).
2. Filtro de escala de grises o escala monocromática.
3. Ecuilizado del histograma.
4. Extracción de características mediante Fisherfaces.
5. Estimación de edad usando el algoritmo del vecino más cercano.

Este proceso está definido en la figura 28.



**Figura 28 Proceso de clasificación de una nueva imagen**

## **Resumen capítulo 3**

En este capítulo se explica la metodología del clasificador la cual esta conformada por la adquisición de las imágenes, el preprocesamiento, la creación de los modelos de representación de características y por último la estimación de edad.

Bajo la metodología dicha anteriormente en el primer punto consistió en investigar en la literatura bases de imágenes que contengan imágenes faciales etiquetadas con información necesaria para obtener la edad de la persona que aparece en la imagen.

En el segundo punto, en el primer paso se realiza el preprocesamiento que consiste en encontrar el rostro de la persona usando clasificadores en cascada, en el segundo paso se utilizan los filtros escala de grises y el ecualizado del histograma obteniendo un conjunto de imágenes estandarizadas.

Para la creación de los modelos de representación de características se utilizó la técnica Fisherface que se explicó en el capítulo 2.

Por último se utiliza el mismo preprocesamiento y la misma técnica de representación de características en una nueva imagen y utilizando la técnica de clasificación por vecindad y se obtiene un resultado de mayoría o minoría de edad.

# Capítulo 4

## Pruebas

Para realizar los experimentos se utilizó una computadora personal con las siguientes características:

Procesador	Intel Celeron N2830 2.16 GHz
Memoria Ram	4GB DDR3
Disco duro	1000GB
Tarjeta de video	Intel® HD Graphics

**Tabla 1. Características de la computadora de pruebas.**

Los experimentos se dividieron en dos partes

1. Clasificación de mayoría y minoría de edad.
2. Estimación de edad

En los cuales se buscó conocer si la cantidad de elementos con los que se entrenan los clasificadores influye directamente en el porcentaje de exactitud y acierto.

## Resultados

### Clasificación de mayoría y minoría de edad.

Un clasificador es considerado una variable explicativa que verifica la presencia o ausencia de una condición.

La información que suministra un clasificador, depende de su nivel de medición: nominal si sus valores son categorías, binaria si es nominal con solo dos categorías, ordinal si las categorías

pueden ser ordenadas de alguna manera, discreta si se puede obtener un número finito de posibles valores distintos y continua si puede asignarse cualquier valor en un rango finito o infinito de valores [67].

En situaciones de clasificación en las que se usan solo dos clases, cada objeto, sujeto o caso es etiquetado con uno de los elementos del conjunto {"positivo (P)", "negativo (N)"} asegurándose la clase a la que pertenece cada caso. En caso concreto nuestro clasificador está usando dos clases; Mayores y menores de edad y el detector de pedofilia solo actuará cuando se detecte un menor de edad en una imagen con contenido pornográfico por lo tanto se etiqueta a la clase menores de edad como "positivos" y a la clase mayores de edad como "negativos".

La aplicación de un clasificador en situaciones como la descrita, produce cuatro posibles resultados:

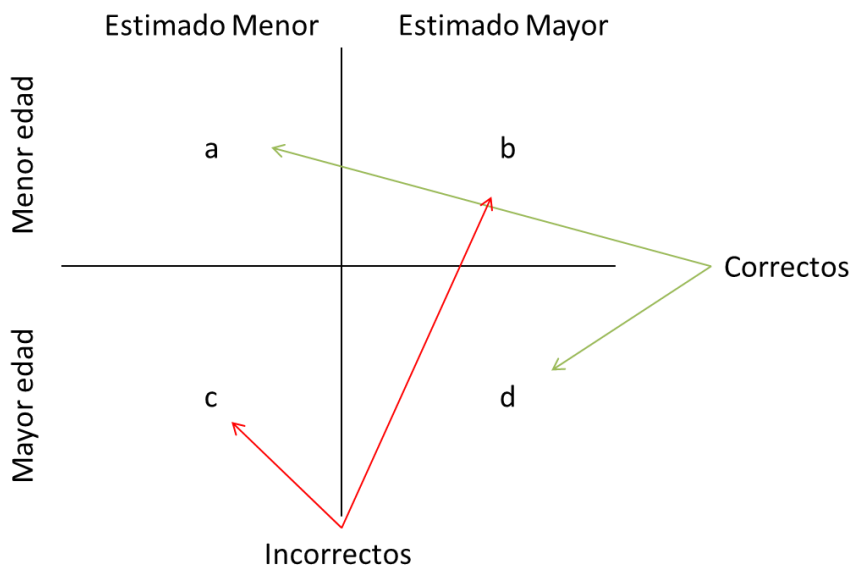
1. Si el caso es positivo y es clasificado como positivo se cuenta como un *verdadero positivo*.
2. Si el caso es positivo y es clasificado como negativo se cuenta como un *falso negativo*.
3. Si el caso es negativo y es clasificado como negativo se cuenta como un *verdadero negativo*.
4. Si el caso es negativo y es clasificado como positivo se cuenta como un *falso positivo*.

Dado un clasificador y un conjunto de sujetos, una matriz de confusión o tabla de contingencia puede ser construida para representar la disposición de dicho conjunto figura 29.

	Estimado Menor	Estimado Mayor
Menor edad	Verdaderos positivos	Falsos positivos
Mayor edad	Falsos negativos	Verdaderos negativos

**Imagen 29. Matriz de confusión.**

La exactitud que tiene un clasificador se mide mediante el número de aciertos que obtiene en función del número de pruebas que se llevan a cabo, esta puede ser interpretada mediante la matriz de confusión donde los números que se encuentran a lo largo de la diagonal principal representan las clasificaciones correctas y los que están a lo largo de la diagonal secundaria representan los errores entre las clases como se observa en la figura 30.



**Figura 30. Interpretación de valores en la matriz de confusión.**

Donde la exactitud del sistema está dada por la ecuación 23.

$$E = \frac{a + d}{a + b + c + d} \quad (22)$$

Sin embargo, un sistema que acierte siempre puede asignar gran cantidad de falsos positivos, y por el contrario, un clasificador estricto que minimice los falsos positivos puede llegar a aumentar el número de falsos negativos. Es por ello que la evaluación de un clasificador considera los aspectos anteriores y proporciona una medida que evalúa simultáneamente la capacidad de discernir a que clase pertenece cada dato. Por lo tanto, se utilizan tres criterios basados en la matriz de confusión mostrada en la figura 29 que son:

***Tasa de aciertos o Recall ecuación 24:***

$$\text{Recall} = \frac{a}{a + b} \quad (23)$$

***Precisión ecuación 25:***

$$\text{Precision} = \frac{a}{a + c} \quad (24)$$

***Factor F ecuación 26:***

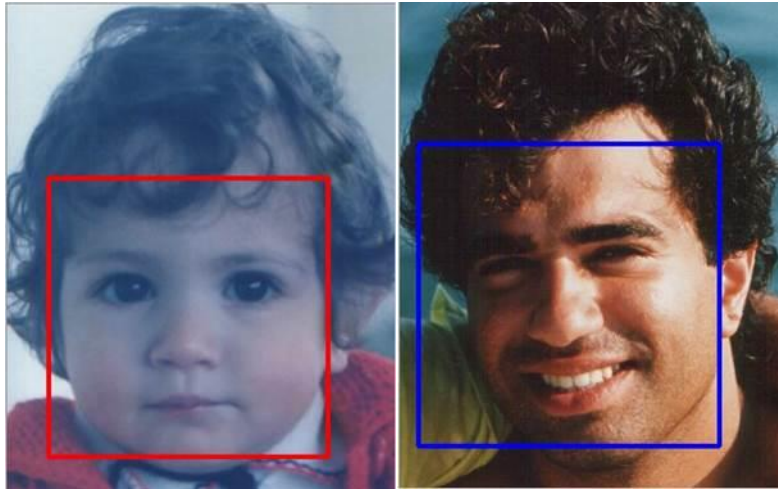
$$F = \frac{2 * (\text{Precision} * \text{Recall})}{\text{Precision} + \text{Recall}} \quad (25)$$

Para la primera parte del experimento se utilizaron las imágenes de entrenamiento de la base de datos FG-Net e IMBD-Wiki generando dos clases representativas, mayores (+18) y menores (-18) de edad obteniendo los resultados mostrados en la tabla 2.

Total de Imágenes	Exactitud	Precisión	Recall	F
10	43%	0.79	0.46	0.58
50	60%	0.84	0.62	0.71
90	64%	0.86	0.68	0.73
120	70%	0.87	0.76	0.81
150	71%	0.89	0.73	0.81
500	74%	0.91	0.76	0.81
750	77%	0.90	0.82	0.84
1000	81%	0.94	0.84	0.85
1500	83%	0.95	0.88	0.85
2000	89%	0.97	0.89	0.86

**Tabla 2. Resultados con la base de datos Fg-Net.**

En la figura 31 se observa la salida visual del clasificador donde el rectángulo en color rojo indica que el rostro en la imagen es menor de edad y el rectángulo color azul indica que es un mayor de edad.



**Figura 31. Clasificación visual de mayoría y minoría de edad.**

Comparando los nuestros resultados con experimentos similares se obtienen los resultados mostrados en la tabla 3:

Fisherfaces (Propio)	Modelo antropométrico [20]	AMM [24]	Eigenfaces [40]	BIF [45]
89%	78%	86.3%	72.2%	73.1%

**Tabla 3. Comparación con trabajos relacionados en clasificación de mayoría y minoría de edad.**

### **Estimación de edad.**

Para la segunda parte del experimento se utilizaron las imágenes que proporciona IMDB-Wiki, para construir los modelos que permiten al clasificador estimar la edad en las imágenes faciales generando una clase por edad (0-90 años) con la misma cantidad de imágenes por clase.

Cuando se trata la edad como un conjunto de valores secuenciales, se puede evaluar el clasificador como un problema de regresión donde el error medio absoluto (MAE) es la manera estándar de expresar los resultados para los sistemas que realizan estimación de edad y se calcula mediante la ecuación 27.

$$\sum_{k=1}^N \frac{|\hat{l}_k - l_k|}{N} \quad (26)$$

Donde  $l_k$  es la edad verdadera de la imagen de prueba  $k$ ,  $\hat{l}_k$  es la edad estimada y  $N$  es el número total de imágenes [26].

Para la segunda parte del experimento se utilizaron las imágenes de la base de datos IMBD-Wiki generando una clase representativa para cada edad en un rango de 0 a 90 años utilizando en cada clase una cantidad igual de imágenes para el modelado de cada clase, obteniendo los resultados mostrados en la tabla 4.

Imágenes por edad	Imágenes totales de entrenamiento	MAE
5	450	7.04
10	900	6.24
15	1350	6.04
20	1800	5.05

**Tabla 4. Resultados estimación de edad.**

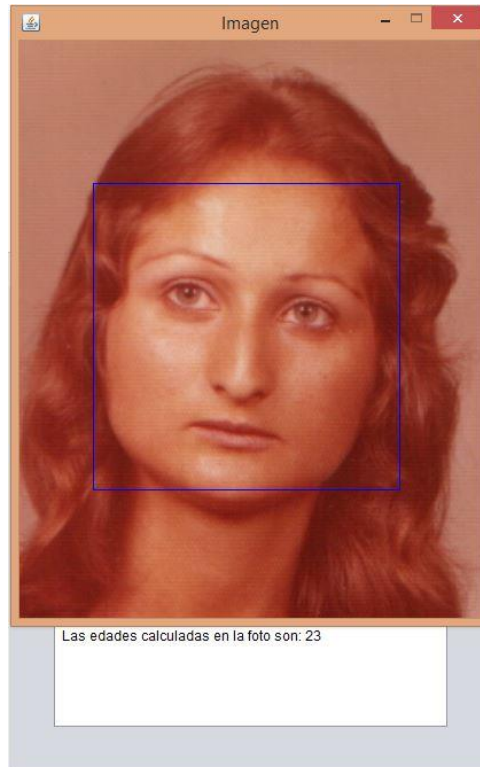
Comparando los nuestros resultados con experimentos similares se obtienen los resultados de la tabla 5:

Fisherfaces (IMBD -Wiki)	BIF (YGA) [57]	LARR(FG-net) [30]	AMM(FG-net) [40]
5.05	3.69	5.8	6.9

**Tabla 5 Resultados con trabajos similares en estimación de edad**

El valor de MAE que obtienen los sistemas, además de estar ligados a la calidad de técnica usada para estimar la edad, también está relacionado con la base de datos utilizada en la implementación. Si esta es heterogénea en la distribución de las edades, es decir, las edades están concentradas en torno a un valor, al entrenar y testear con esa base de datos, lo normal es que se obtengan valores de MAE pequeños utilizar un margen de edades pequeño como los realizados por [68], [57] y [40] por lo que un MAE de bajo valor indica que el sistema funciona satisfactoriamente [68] por ultimo en los sistemas clasificadores de edad lo deseable es tener un valor de MAE lo más cercano a cero como sea posible. Sin embargo, se sabe que esto no es posible ya que el proceso de envejecimiento es algo personal e incontrolable.

En la figura 32 se observa el resultado visual del estimador de edad indicando una edad de 23 años y cuya edad real es de 19 años.



**Figura 32. Estimación visual de edad.**

## **Resumen capítulo 4**

En el capítulo número cuatro se explica la matriz de confusión la cual es una herramienta que permite medir la exactitud de un clasificador.

Además se muestran los resultados obtenidos en las pruebas realizadas en la clasificación de la mayoría y minoría de edad y el error medio absoluto al estimar la edad exacta en los rostros de una imagen.

## Conclusiones

La estimación de edad automática de edad, específicamente la clasificación de mayores y menores de edad es un problema que impacta en distintas áreas con múltiples aplicaciones potenciales, por ello en este trabajo de fin de grado se desarrolló un clasificador basado en apariencia utilizando los modelos de representación de características Fisherfaces los cuales son una combinación de los métodos estadísticos análisis de componentes principales y análisis discriminante lineal.

Mediante los resultados conseguidos en esta investigación y en comparación con los previamente investigados se puede concluir que el clasificador obtiene mejores resultados con imágenes provenientes de entornos reales (cámara web, fotogramas de video y fotografías en entornos no controlados) a diferencia de los clasificadores basados en modelos antropométricos y puntos fiduciales los cuales solamente pueden clasificar imágenes donde el rostro se encuentra completamente de frente y en entornos controlados.

También existen investigaciones de clasificadores basados en apariencia que obtienen resultados más precisos debido a que realizan un preprocesamiento más exhaustivo sobre las imágenes de entrenamiento como Kyle Patterson en [40], que realiza un etiquetado de las imágenes por grupo etario y sexo y Kwon Young en [20] el cual obtiene todos los elementos característicos del rostro (ojos, nariz y boca) lo cual representa un alto costo computacional.

Para la detección del rostro, se utilizó el filtro en cascada propuesto por Viola and Jones [59] que está incluido en la biblioteca opencv debido a que es más rápido que y el ratio de detección es ligeramente inferior que otros métodos.

En el apéndice A se muestran los experimentos realizados al clasificador con distintos filtros demostrando que se obtienen mejores resultados al aplicar un filtro a escala de grises y ecualizar el histograma a cuando se utiliza un solo filtro o en su defecto no se utilizan filtros.

Se demostró que el porcentaje de aciertos es proporcional al número de imágenes utilizadas para la creación de los modelos como lo afirman Santos et al. En [57] y Suárez Ricardo en [65], debido a que se extraen más características representativas de cada clase, indicando que si se tiene mayor número de imágenes de entrenamiento se obtienen mejores resultados en la estimación. Esto se demuestra al verificar que los mejores resultados se consiguieron cuando se utilizaron 2000 imágenes para la clasificación de mayores y menores de edad obteniendo un 89% de exactitud y 1800 imágenes en la estimación de edad obteniendo un error medio absoluto de 5.05

Por último, se observó que el número de imágenes de entrenamiento utilizados influye directamente sobre el tamaño y el tiempo de creación de los modelos de representación de características.

El trabajo a futuro corresponde a generar una metodología basada en multiprocesamiento con GPU que permita trabajar con grandes volúmenes de imágenes sin que esto conlleve un problema de rendimiento.

Por último se recomienda el uso de la base de datos FG-net la cual es ampliamente usada en investigaciones como las realizadas por Luu et al. En [58], Guo et al. En [56] y Hewahi et al. En [47], debido a que proporciona un conjunto de imágenes de entrenamiento y de prueba con su respectiva edad.

## Referencias

1. Jiménez, A. G. (2011). Una perspectiva sobre los riesgos y usos de internet en la adolescencia. *Revista ICONO14*, 9(3), 410-425.
2. García-Piña, C. A. (2008). Riesgos del uso de internet por niños y adolescentes. *Estrategias de seguridad. Acta Pediátrica de México*, 29(5), 272-278.
3. Castro, O. J. Pornografía, pedofilia y control de contenidos en Internet.
4. Acosta Rojas, M. D. P., & Acosta Silva, J. (2009). McDonald's: un ejemplo de publicidad estratégica dirigida a los niños (Bachelor's thesis).
5. Panasonic.com. (2017). PANASONIC SMART TV LA TV QUE TE GUSTA SABE LO QUE TE GUSTA - Panasonic España. [online] Available at: <https://www.panasonic.com/es/consumer/televisores-aprender/televisores/panasonic-smart-tv-recomendacion-my-stream.html> [Accessed 1 Sep. 2017].
6. B. BREEDEN AND J. MULHOLLAND, "INVESTIGATING 'INTERNET CRIMES AGAINST CHILDREN' (ICAC) CASES IN THE STATE OF FLORIDA," PRESENTED AT THE PROCEEDINGS OF THE 2006 ACM SYMPOSIUM ON APPLIED COMPUTING, DIJON, FRANCE, 2006.
7. P. F. A. A. PATTERSON, "PAEDOPHILE INTERNET ACTIVITY," AUSTRALIAN INSTITUTE OF CRIMINOLOGY.
8. L. PENNA, ET AL., "CHALLENGES OF AUTOMATING THE DETECTION OF PAEDOPHILE ACTIVITY ON THE INTERNET," IN PROC. FIRST INTERNATIONAL WORKSHOP ON SYSTEMATIC APPROACHES TO DIGITAL FORENSIC ENGINEERING, 2005, PP. 206--220.] y [K. F. DURKIN, "MISUSE OF THE INTERNET BY PEDOPHILES: IMPLICATION FOR LAW ENFORCEMENT AND PROBATION PRACTICE.," FEDERAL PROBATION, VOL. 61, PP. 14-18, 1997
9. Carr, J. y Z. Hilton, *Digital Manifesto*, Children's Charities Coalition on Internet Safety, Londres, 2009, pág. 29.
10. Internet Watch Foundation, *Annual and Charity Report 2010*, IWF, Londres, 2010, pág. 1; Quayle, E. y T. Jones, "Sexualized images of children on the internet-,"
11. Internet Watch Foundation, *Annual and Charity Report 2010*, IWF, Londres, 2010, disponible en <[www.iwf.org.uk/accountability/annualreports/2010-annual-report](http://www.iwf.org.uk/accountability/annualreports/2010-annual-report)>;
12. Internet Watch Foundation, *Annual and Charity Report 2006*, IWF, Londres, 2006, pág. 8, [www.enough.org/objects/20070412\\_iwf\\_annual\\_report\\_2006\\_web.pdf](http://www.enough.org/objects/20070412_iwf_annual_report_2006_web.pdf)
13. Internet Watch Foundation, *Annual and Charity Report 2010*, pág. 8
14. Steiner, M. (2010). Facial Image-based Age Estimation. Study Thesis at the Faculty of Computer Science Institute for Anthropomatics Facial Image Processing and Analysis.
15. Alegre, E., de Paz, I., Gonzales, V., Garcia, O., Teresa, M. and Fidalgo, E. (2016). Detección de rostro y estimación de edad y género utilizando sistemas de visión portables de bajo coste. In: *Cátedra Telefónica*. [online] Leon, pp.21-46. Disponible

- en:[http://catedratelefonica.unileon.es/wpcontent/uploads/2016/05/cuaderno\\_catedra\\_ule\\_vol\\_3\\_2016.pdf](http://catedratelefonica.unileon.es/wpcontent/uploads/2016/05/cuaderno_catedra_ule_vol_3_2016.pdf) [Accessed 4 Sep. 2017].
16. Fu, Y., Guo, G., & Huang, T. S. (2010). Age synthesis and estimation via faces: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 32(11), 1955-1976.
  17. Alcohol servers' estimates of young people's ages. Willner, P. and Rowe, G. 2001, *Drugs: Education, Prevention, and Policy*, 8, pp. 375–383.
  18. Selling tobacco to minors: Can merchants accurately determine a customers age? Jason, L. A., Pokorny, S. B., Sherk, J. L., Helzing, M. D. y Rebus, P. J. 2003, *Journal of Human Behavior in the Social Environment*, 8, pp. 67-73.
  19. F. Yun, G. Guodong, and T. S. Huang, "Age Synthesis and Estimation via Faces: A Survey," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, pp. 1955-1976, 2010.
  20. Y. Kwon and N. Lobo, "Age Classification from Facial Images," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 762-767, 1994.
  21. N. Ramanathan and R. Chellappa, "Modeling Age Progression in Young Faces," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 387-394, 2006.
  22. Farkas, L. G., & Deutsch, C. K. (1996). Anthropometric determination of craniofacial morphology. *American Journal of Medical Genetics Part A*, 65(1), 1-4.
  23. Y. Kwon and N. Lobo, "Age Classification from Facial Images," *Computer Vision and Image Understanding*, vol. 74, no. 1, pp. 1-21, 1999.
  24. Cootes, T. F., Edwards, G. J., & Taylor, C. J. (2001). Active appearance models. *IEEE Transactions on pattern analysis and machine intelligence*, 23(6), 681-685.
  25. Lanitis, A., Taylor, C. J., & Cootes, T. F. (2002). Toward automatic simulation of aging effects on face images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(4), 442-455.
  26. Geng, X., Zhou, Z. H., & Smith-Miles, K. (2007). Automatic age estimation based on facial aging patterns. *IEEE Transactions on pattern analysis and machine intelligence*, 29(12), 2234-2240.
  27. Image-Based Human Age Estimation by Manifold Learning and Locally Adjusted Robust Regression. Guo, G., Fu, Y., Dyer, C.R. y Huang, T.S.7, Julio 2008, *IEEE Transactions on Image Processing*, Vol. 17.
  28. Cai, D., He, X., Han, J., & Zhang, H. J. (2006). Orthogonal laplacianfaces for face recognition. *IEEE transactions on image processing*, 15(11), 3608-3614.
  29. Fu, Y., & Huang, T. S. (2008). Human age estimation with regression on discriminative aging manifold. *IEEE Transactions on Multimedia*, 10(4), 578-584.
  30. Guo, G., Fu, Y., Dyer, C. R., & Huang, T. S. (2008). Image-based human age estimation by manifold learning and locally adjusted robust regression. *IEEE Transactions on Image Processing*, 17(7), 1178-1188.

31. Yan, S., Wang, H., Fu, Y., Yan, J., Tang, X., & Huang, T. S. (2009). Synchronized submanifold embedding for person-independent pose estimation and beyond. *IEEE Transactions on Image Processing*, 18(1), 202-210.
32. Ahonen, T., Hadid, A., & Pietikainen, M. (2006). Face description with local binary patterns: Application to face recognition. *IEEE transactions on pattern analysis and machine intelligence*, 28(12), 2037-2041.
33. Sanhueza R.: Análisis y detección de características faciales usando aprendizaje estadístico. Tesis del Universidad de Chile, Facultad de ciencias Físicas y Matemáticas, (2008).
34. Hayashi J., Yasumoto M., Ito H., & Koshimizu H.: Method for estimating and modeling age and gender using facial image processing in *Virtual Systems and Multimedia*, pp. 439-448, *IEEE Proceedings. Seventh International Conference* (2001).
35. Suo, J., Min, F., Zhu, S., Shan, S., & Chen, X. (2007, June). A multi-resolution dynamic model for face aging simulation. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on* (pp. 1-8). IEEE.
36. Belhumeur P., Hespanha J. & Kriegman D.: Eigenfaces vs fisherfaces: Recognition using class specific linear projection. *Transactions on pattern analysis and machine intelligence* 19(7), pp. 711-720, (1997).
37. Terradez G.: Análisis de componentes principales. *Universidad Oberta de Cataluña* 11 (2002)
38. Hernandez J. & Mendez N.: Reconocimiento de expresiones faciales para interacción con el computador. *Jovenes en la ciencia* 2(1) 62-66 (2016).
39. Belhumeur P., Hespanha J. & Kriegman D.: Eigenfaces vs fisherfaces: Recognition using class specific linear projection in *Bernard F. Buston and Roberto Cipolla*, pp: 45-58, vol 1064, Springer (1996).
40. Patterson, K. Automatic Age Estimation and Interactive Museum Exhibits. *Experimental Aging Research*, 35(2), 268-275, (2014).
41. F. Yun, G. Guodong, and T. S. Huang, "Age Synthesis and Estimation via Faces: A Survey," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, pp. 1955-1976, 2010.
42. Dasarathy, B. V. *Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques*. 1991.
43. Tin H.: Gender and Age Estimation Based on Facial Images. *Acta technical Napocensis* 52(3) 37 (2011)
44. Betancourt, G. A. (2005). Las máquinas de soporte vectorial (svms). *Scientia et technica*, 1(27).
45. Guo, G., Mu, G., Fu, Y., & Huang, T. S. (2009, June). Human age estimation using bio-inspired features. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on* (pp. 112-119). IEEE.

46. Ramesha K, K B Raja, Venugopal K R, L M Patnaik: Feature Extraction based Face Recognition, Gender and Age Classification. (IJCSSE) International journal on Computer Science and Engineering, Vol. 02, No.01S, 2010, 14-23.
47. Hewahi N., Olwan a., Tubeel N., El-Asar S. & Abusultan Z.: Age estimation base don neural networks using face features. Journal of Emerging Trends in Computing and Information Sciences, 1(2). 61-67 (2010).
48. Lanitis, A., Taylor, C. J., & Cootes, T. F. (2002). Toward automatic simulation of aging effects on face images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(4), 442-455.
49. Fu, Y., & Huang, T. S. (2008). Human age estimation with regression on discriminative aging manifold. IEEE Transactions on Multimedia, 10(4), 578-584.
50. Guareño, M., & José, J. (2016). Support vector regression: propiedades y aplicaciones.
51. Zhou, S. K., Georgescu, B., Zhou, X. S., & Comaniciu, D. (2005, October). Image based regression using boosting method. In Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on (Vol. 1, pp. 541-548). IEEE.
52. Image-Based Human Age Estimation by Manifold Learning and Locally Adjusted Robust Regression. Guo, G., Fu, Y., Dyer, C.R. y Huang, T.S.7, Julio 2008, IEEE Transactions on Image Processing, Vol. 17.
53. Intel Corporation. The Open Source Computer Vision (OpenCV) Library Homepage. URL: <http://opencv.org/>
54. Rothe R., Timofte R. & Van G.: Dex: Deep proceeding in Proceedings of the IEEE. International Conference on Computer Vision Workshops, pp: 10-15, (2015).
55. FG-Net available: <http://www.fgnet.rsunit.com/>
56. Gou G., Gu Y., Dyer C. & Huang T.: Image-based human age estimation by manifold learning and locally adjusted robust regression. IEEE Transactions on Image Processing 17(7), 1178-1188, (2008).
57. Guo, G., Fu, Y., Dyer, C. R., & Huang, T. S. (2008). Image-based human age estimation by manifold learning and locally adjusted robust regression. IEEE Transactions on Image Processing, 17(7), 1178-1188.
58. Luu K., Ricanek K., Bui T. & Suen C.: Age estimation using active appearance models and support vector machine regression in Biometrics: Theory, Applications and Systems. IEEE 3rd International Conference, pp: 1-6, (2009).
59. P. Viola and M.J. Jones. Rapid object detection using a boosted cascade of simple features. In IEEE Intl. Conf. on Computer Vision and Pattern Recognition, CVPR 2001, pages 12–14, Kauai, Hawaii, 2001.
60. Intel Corporation. The Open Source Computer Vision (OpenCV) Library Homepage. URL: <http://opencv.org/>

61. M.-H. Yang, D. Roth, and N. Ahuja. A SNoW-based face detector. In S.A. Solla, T. K. Leen, and K.-R. Muller, editors, *Advances in Neural Information Processing Systems*, pages 855–861. MIT Press, 2000.
62. H.A. Rowley. *Neural Network-Based Face Detection*. PhD thesis, Carnegie Mellon University, 1999.
63. Esquivel, H., & Natanael, G. (2017). Sistema embebido para la segmentación de imágenes en tiempo real.
64. Barrero, A., Robayo, M., & Jacinto, E. (2015). Algoritmo de navegación a bordo en ambientes controlados a partir de procesamiento de imágenes. *Tekhnê*, 12(2), 23-34.
65. Suárez Fuentes: Diferenciación visual de niños y adultos a partir de imágenes faciales (2013).
66. Cox I., Ghosn J. & Yianilos P.: Features based face recognition using mixture-distance in *Computer Vision and Pattern Recognition*, pp. 209-216, IEEE computer Society Conference (1996).
67. Noguera, T. (2010). Metodología ROC en la evaluación de medidas antropométricas como marcadores de la hipertensión arterial. Aplicación a población gallega adulta (tesis). Santiago de Compostela: Universidad Santiago de Compostela.
68. Rodríguez Ruiz, A. (2012). Estimulación automática de edad a partir de imágenes faciales (Bachelor's thesis).

## Apéndice A Comparación de filtros

En esta sección se muestran los experimentos realizados para demostrar el aumento de exactitud en la clasificación de mayoría y minoría de edad con la combinación de los filtros; Escala de grises y ecualizado del histograma.

Se utilizaron los elementos de prueba de la base de imágenes Fg-Net la cual consta de 167 elementos donde 143 son menores de edad y 24 son mayores de edad y el método de evaluación fue una matriz de confusión y curvas ROC.

En las siguientes imágenes se observa el resultado del clasificador final utilizando la técnica Fisherfaces, en la figura A1 se utilizaron 45 imágenes por clases y no se aplicó ningún filtro sobre las imágenes.

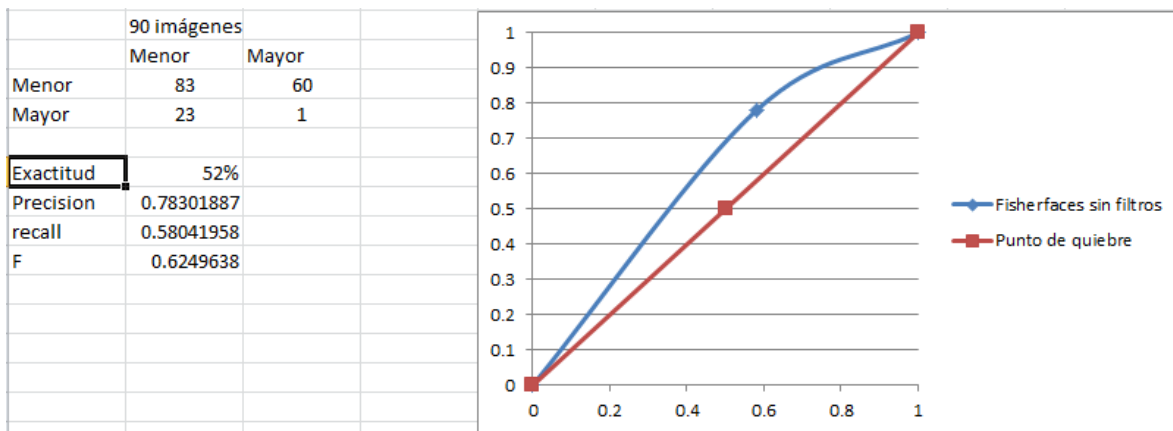
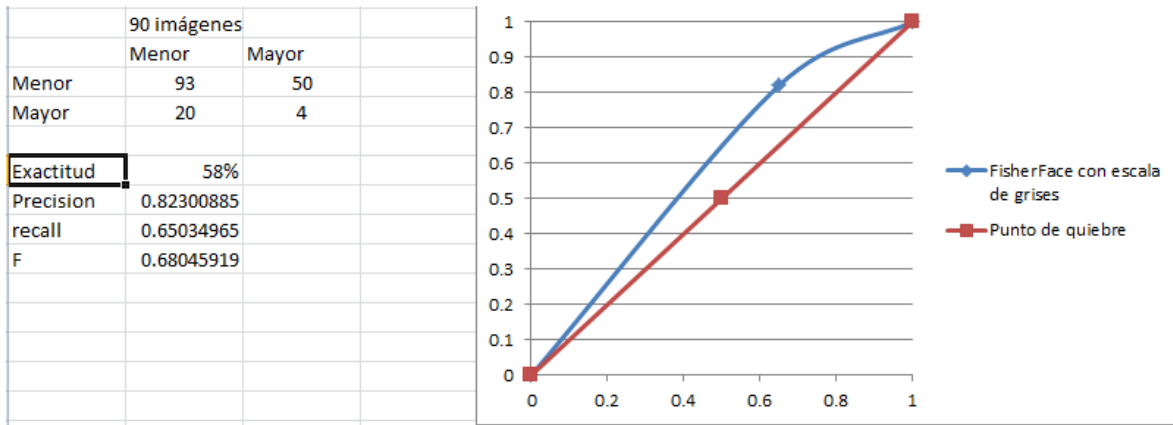


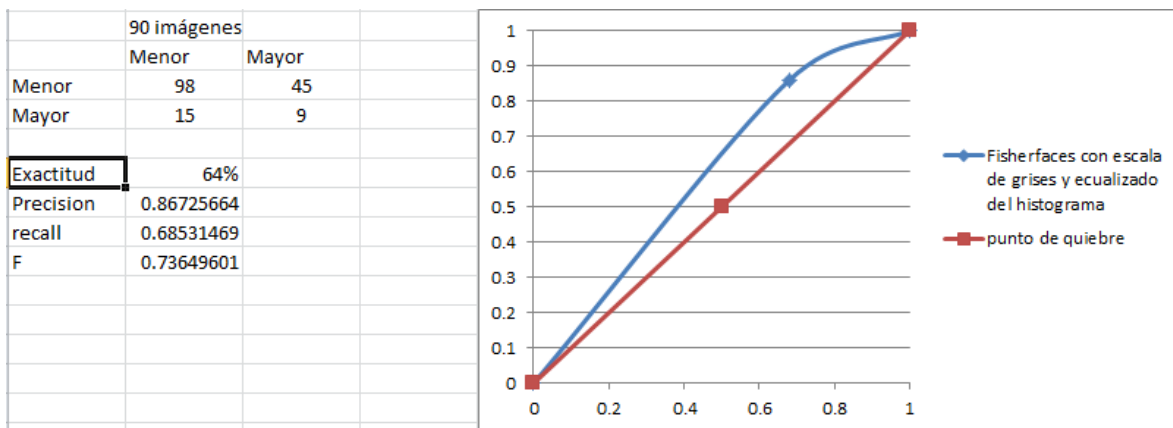
Figura A1. Resultados sin filtros.

En la figura A2 se utilizaron con 45 imágenes por clase con escala de grises.



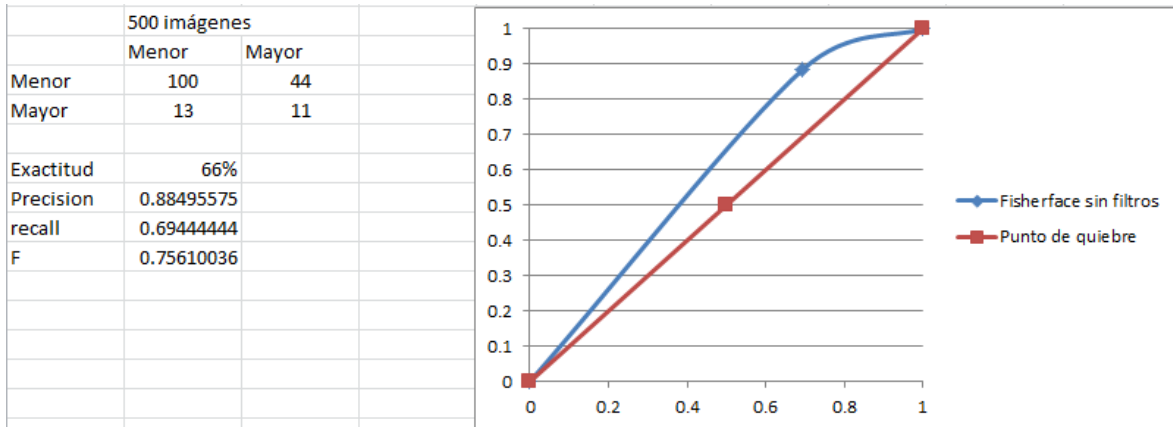
**Figura A2. Resultados con escala de grises.**

En la figura A3 se utilizaron con 45 imágenes por clase con escala de grises y el ecualizado del histograma, mostrando un incremento en la exactitud del clasificador.



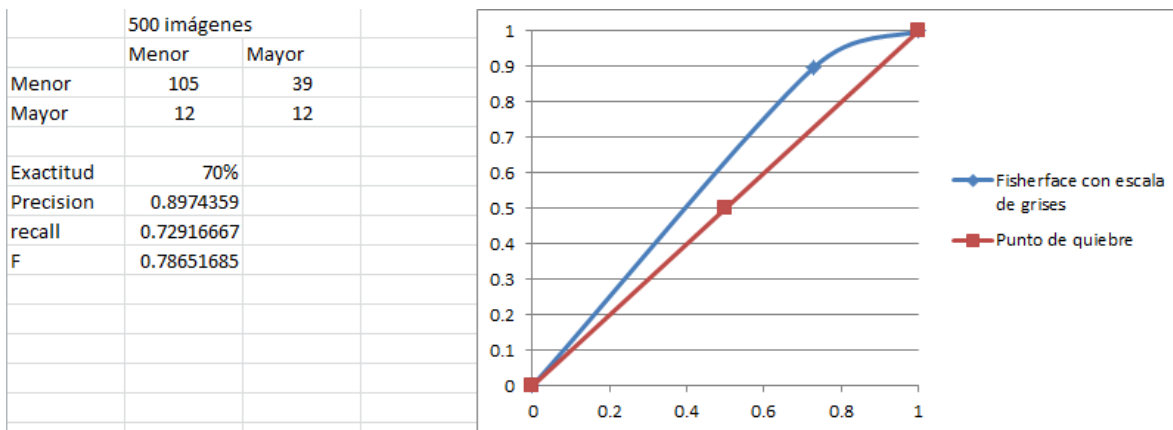
**Figura A3. Resultados con escala de grises y ecualizado de histograma.**

En la figura A4 a la figura A6 se utilizaron 250 imágenes por clase, en la figura A4 no se utilizó ningún filtro.



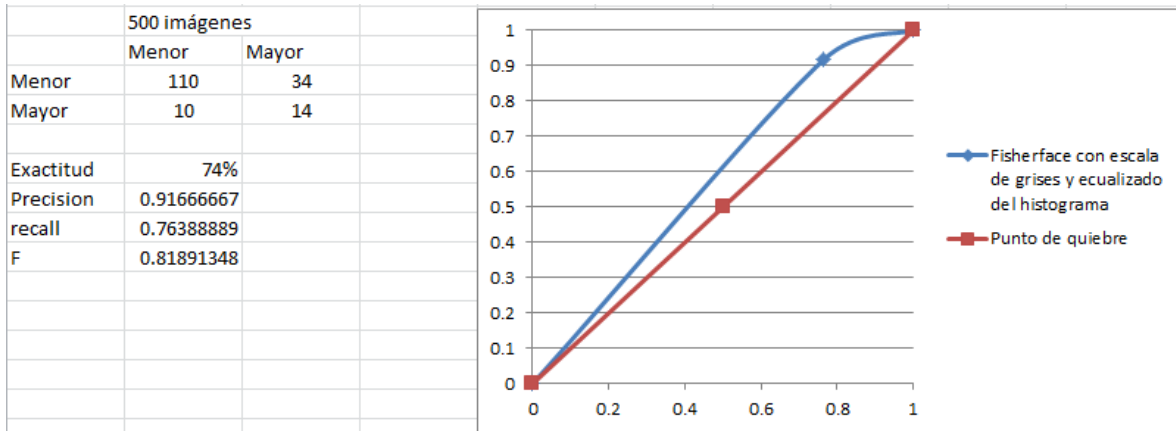
**Figura A4. Resultados sin filtros.**

En la figura A5 se utilizó el filtro escala de grises.



**Figura A5. Resultados con escala de grises.**

En la figura A6 se utilizó el filtro escala de grises y el ecualizado del histograma, mostrando nuevamente el incremento en la exactitud del clasificador.



**Figura A6. Resultados con escala de grises y ecualizado del histograma.**