

**Prototipo de sistema de control parental basado en
estimación automática de edad extraída de imágenes
faciales**



Wilian Fernando Pantoja Carvajal
William Alexis Ruano Daza

Directora: Ph.D. (c) Maritza Fernanda Mera Gaona
Codirectora: Ph.D. (c) Sandra Milena Roa Martínez

Monografía

Universidad del Cauca

Facultad de Ingeniería Electrónica y Telecomunicaciones
Departamento de Sistemas
Grupo de Investigación en Inteligencia Computacional
Línea de investigación en Reconocimiento de Patrones
Popayán, diciembre de 2017

Tabla de contenido

1. INTRODUCCIÓN.....	7
1.1. Planteamiento del problema	7
1.2. Justificación.....	9
1.3. Objetivos	10
1.3.1. Objetivo general	10
1.3.2. Objetivos específicos.....	11
1.4. Aportes.....	11
1.4.1. Investigativos.....	11
1.4.2. De Innovación	11
1.5. Estructura del documento.....	12
2. MARCO TEÓRICO Y ANTECEDENTES	13
2.1. Marco teórico.....	13
2.1.1. Visión por computador.....	13
2.1.2. Representación de una imagen.....	15
2.1.3. Descriptores de texturas.....	18
2.1.4. Sistema de reconocimiento de patrones.....	19
2.1.5. Máquinas de soporte vectorial.....	21
2.1.6. Estimación automática de edad.....	23
2.2. Antecedentes	24
2.2.1. Modelos de representación facial	24
2.2.2. Sistemas de Estimación Automática de Edad.....	32
2.2.3. Sistemas de Control Parental	34
2.2.4. Sistemas de Control de Acceso y Seguridad.....	36
2.3. Conclusiones Parciales	37
3. MECANISMO PROPUESTO DE ESTIMACIÓN DE LA EDAD.....	39
3.1. Detección y Normalización	40
3.1.1. Detección facial	40
3.1.2. Detección de puntos de referencia (<i>landmarks</i>).....	42
3.1.3. Normalización.....	44
3.2. Extracción de características	46
3.3. Clasificación	55

3.4.	Diagrama funcional del mecanismo propuesto de estimación de la edad	56
4.	EVALUACIÓN DEL MECANISMO DE ESTIMACION DE LA EDAD PROPUESTO ..	58
4.1.	Protocolo de Evaluación Comparativa	58
4.1.1.	Base de datos	58
4.1.2.	Definición del conjunto de entrenamiento y prueba	59
4.1.3.	Medición de Precisión.....	59
4.2.	Implementación del protocolo	60
5.	PROTOTIPO DE CONTROL PARENTAL.....	68
5.1.	Requerimientos	68
5.1.1.	Otras consideraciones	69
5.2.	Análisis y Diseño	69
5.2.1.	Definición de actores	69
5.2.2.	Casos de uso	70
5.2.3.	Arquitectura del sistema	71
5.3.	Implementación del prototipo.....	74
5.3.1.	Vista de implementación.....	74
5.3.2.	Proceso de desarrollo.....	77
5.4.	Implantación del prototipo.....	86
5.4.1.	Vista de despliegue	86
6.	CONCLUSIONES Y TRABAJO FUTURO.....	88
6.1.	Conclusiones.....	88
6.2.	Trabajo futuro	89
7.	REFERENCIAS BIBLIOGRÁFICAS.....	90

Índice de figuras

Figura 1. Histograma de una imagen. a) Imagen original $f(x, y)$, b) componentes de color de la imagen original RGB, rojo, verde y azul de arriba abajo. c) Imagen en escala de grises $g(x, y)$. d) Histograma de la imagen, la línea en blanco representa el histograma para la imagen en escala de grises.	16
Figura 2. Ampliación de una imagen y vecindario de un pixel.	17
Figura 3. Coordenadas de los vecinos de un pixel p	18
Figura 4. Diagrama de bloques de un modelo estadístico de reconocimiento de patrones.	19
Figura 5. Hiperplano óptimo entre dos clases (azul y rojo) de muestras. Imagen tomada de [34].	21
Figura 6. Proceso de un sistema de estimación de edad.	23
Figura 7. Vista general del mecanismo de estimación de edad.	39
Figura 8. Vista general de la etapa de detección y normalización.	40
Figura 9. Tres window sliding y sus respectivos puntajes.	41
Figura 10. Ejemplo de aplicación de método de detección de puntos de referencia (a) Características ejemplo. (b) Resultado ideal. (c) Resultado al aplicar LNF (Adaptado de CLNF).	44
Figura 11. Ejemplo del proceso completo de detección y normalización. (a) Imagen de entrada. (b) Detección facial. (c) Detección de landmarks. (d) Imagen normalizada.	45
Figura 12. Vista general de la etapa de extracción y selección de características.	46
Figura 13. Ejemplo del Operador LBP básico.	47
Figura 14. Generación de patrón binario inconsistente en el proceso de codificación LBP.	47
Figura 15. Operador CLBP y cálculo de los sub-CLBP patrones.	48
Figura 16. Proceso de extracción de características con CLBP. (a) División en celdas de la imagen. (b) Extracción de los códigos CLBP. (c) Histograma celda a celda. (d) Histograma general de la imagen.	49
Figura 17. Calculo de gradiente.	50
Figura 18. Proceso extracción HOG.	51
Figura 19. Calculo del histograma de gradientes.	51
Figura 20. Vectores de Gradiente con orientaciones similares asignados a intervalos de orientación diferentes.	52
Figura 21. Pixeles cercanos asignados a celdas diferentes.	53
Figura 22. (a) Definición de los bloques de celdas (b) Contribución de una celda al vecindario de bloques, normalización y vector de características.	54
Figura 23. Vista general de la etapa de clasificación.	55
Figura 24. Diseño del mecanismo de estimación de edad.	57
Figura 25. Ejemplo de imágenes de la base de datos loG.	59
Figura 26. Diagrama de casos de uso para el escenario de configuración y monitoreo ...	70

Figura 27. Diagrama de casos de uso para escenario de servicio de filtrado de paquetes de red	71
Figura 28. Arquitectura del Prototipo.....	73
Figura 29. Diagrama de paquetes de software	74
Figura 30. Diagrama de componentes del componente de estimación de edad	75
Figura 31. Diagrama de componentes del control de red.....	75
Figura 32. Ejemplo de un archivo contenedor de las URL que pertenecen a una categoría	78
Figura 33. Configuración de reglas para un grupo de edad.	79
Figura 34. Ejemplo del proceso de extracción de características: a) Detección de cara y landmarks. b) Escala de grises y recorte. c) Separación en celdas. d) Aplicación del operador LBP. e) Vectores de gradiente para cálculo de HOG.....	80
Figura 39. Página web que es mostrada cuando se detecta el acceso a una URL del control de acceso (lista negra).....	81
Figura 40. Servicio de prototipo de control parental instalado en la estación de trabajo. .	83
Figura 41. Mecanismo de sincronización para la ejecución de comandos sobre la base de datos.....	85
Figura 42. Nodos desplegados en un computador personal.	87

Índice de tablas

Tabla 1. Áreas y ejemplos de aplicación de visión por computador.	13
Tabla 2. Resumen y comparación de técnicas de estimación de edad sobre la base de datos FG-NET.	29
Tabla 3. Resumen y comparación de técnicas de estimación de edad sobre la base de datos MORPH.	30
Tabla 4. Resumen y comparación de técnicas de estimación de edad en otras bases de datos de imágenes faciales.	31
Tabla 5. Resumen y comparación de técnicas de estimación de edad en la base de datos IOG.	32
Tabla 6. Funcionalidades de los sistemas de control parental para dispositivos móviles disponibles en el mercado.	34
Tabla 7. Funcionalidades de los sistemas de control parental para computadores personales disponibles en el mercado.	35
Tabla 8. AEM y AEO obtenidos para cada uno de los descriptores y diferentes valores de k usando el método kNN.	61
Tabla 9. AEM y AEO obtenidos para cada uno de los descriptores y diferentes valores de γ usando el método RF.	62
Tabla 10. AEM y AEO obtenidos para cada uno de los descriptores usando el método LDA.	63
Tabla 11. AEM y AEO obtenidos para cada uno de los descriptores y diferentes valores de C usando el método SVM.	63
Tabla 12. Matriz de confusión de la evaluación para el descriptor LPB+HOG y el método de clasificación SVM con el <i>kernel</i> RBF y $C = 4$	66
Tabla 13. Comparación del mecanismo propuesto con los resultados más relevantes encontrados en los antecedentes.	66
Tabla 14. Descripción de Actor Administrador.	69
Tabla 15. Descripción de Actor Controlador de Red.	69

1. INTRODUCCIÓN

1.1. Planteamiento del problema

El rostro de una persona contiene una serie de características que describen y transmiten una gran cantidad de información [1]. Algunas de las características son estructurales (color, líneas, manchas, textura, y flacidez de la piel), otras pueden ser dinámicas (movimientos musculares y dilatación de la pupila) o artificiales (peinado, gafas, y maquillaje). A partir de estas características faciales los seres humanos tienen la capacidad innata de interpretar atributos como sexo, raza e identidad; así como también información asociada a emociones, edad y, condición física y mental [2].

Algunos estudios realizados han demostrado que las personas tienen la habilidad de determinar con alta precisión la edad de personas entre los 20 y 60 años. En la mayoría de estas investigaciones se encontró una diferencia de entre 3 y 7 años respecto a la edad real de la persona [2] [3], siendo una de las principales dificultades de la estimación de la edad, los cambios del rostro producidos con el paso del tiempo, debido a los procesos de crecimiento y envejecimiento. Ciertos cambios están asociados al tamaño, estado y posición de los atributos faciales, por ejemplo, la calidad de la piel es uno de los atributos más afectados con el paso del tiempo ya que progresivamente se arruga, se mancha y se incrementa la cantidad de poros abiertos [2]. En este sentido, se encuentran diversas propuestas de mecanismos de estimación automática de edad capaces de simular la capacidad humana mediante la cual se realiza la estimación aproximada de la edad de un rostro o se clasifica el rostro dentro de un rango de edades [1]. Sin embargo, se han evidenciado numerosos retos de investigación debido a la dificultad de representar los cambios faciales que describen el envejecimiento en una persona, estos retos obedecen en gran medida a que los cambios en el rostro están influenciados por diferentes ritmos de envejecimiento asociados a los genes [4] [5], estado de salud, estilo de vida, ambiente de trabajo y efectos de la radiación ultravioleta [3], entre otros. Incluso, los patrones de envejecimiento varían entre hombres y mujeres debido al uso de maquillajes y accesorios [4].

Actualmente existe una preocupación generalizada sobre la influencia negativa que los contenidos disponibles en medios digitales tienen sobre los menores de edad, dada la facilidad de acceso a diferentes tecnologías y dispositivos que permiten la manipulación de dichos contenidos. Ésta facilidad ha permitido que niños y adultos tengan a su disposición una gran variedad de información a través de la televisión, libros, revistas, internet, dispositivos móviles, entre otros [6]. Considerando lo anterior, se han desarrollado diferentes estudios respecto a la exposición de menores a contenidos violentos, y entre los principales hallazgos identificados se evidencia la relación directa entre comportamientos

agresivos [7], un aumento de la insensibilidad hacia las víctimas de la violencia, y el aumento exagerado del temor de ser víctimas de la violencia [8].

Otra de las grandes problemáticas, es la facilidad de acceso por parte de menores de edad a contenidos sexuales a través de los diferentes medios de comunicación, ya que estos podrían generar problemas sobre el desarrollo psicológico de los menores [6]. De igual forma, la Sociedad Pediátrica Canadiense [9] evidencia la influencia de los medios digitales y las tecnologías en los comportamientos de menores de edad, respecto al aprendizaje, sexualidad, nutrición, entre otros.

Por las preocupaciones expresadas anteriormente, el acceso a los contenidos debe tener restricciones de acuerdo a la información inmersa, debido a que contenidos sexuales, violentos y de lenguaje vulgar podrían producir comportamientos negativos en los menores de edad [8]. En este sentido, los medios de comunicación clasifican sus contenidos de acuerdo a las restricciones del contenido y la edad del público apto para recibirlo. Esta clasificación es notable principalmente en videojuegos, cine, televisión e internet y difiere en cada país o región, un ejemplo de esto lo podemos ver en [10] donde cada contenido se clasifica según el público apto para verlo y la región.

En el contexto nacional la televisión pública clasifica su programación conforme al artículo 25 del acuerdo 002 de 30 de junio de 2011, de la siguiente manera: i) infantil, para niños de 0 a 12 años; ii) adolescente, para edades entre 12 y 18 años; iii) familiar, todas las edades; iv) adulto, mayores de 18 años. La importancia de la restricción de acceso a contenidos también se sustenta en el código de infancia y adolescencia, ley número 1090 del 8 de Noviembre de 2006, el cual reza que el Estado debe velar por la integridad de los niños [11]. Adicionalmente, la Convención sobre los Derechos del Niño (CDN)¹ en su Artículo 17 resalta que es deber del Estado promover la elaboración de directrices apropiadas para proteger al niño contra toda información y material perjudicial para su bienestar [12].

En este sentido, se hace necesario controlar el acceso a la información, contenidos, productos, entre otros, para proteger de material inadecuado a menores de edad, sin embargo, los métodos existentes están expuestos a que el usuario pueda engañar al sistema sobre su edad debido a que la configuración y autenticación de estas herramientas es manual. También, algunos autores [4] [5] [13] han propuesto emplear la estimación automática de la edad para soportar la restricción de contenidos asociados a ésta. Y surge

¹ CDN: Es un tratado internacional que recoge los derechos de la infancia y es el primer instrumento jurídicamente vinculante que reconoce a los niños y niñas como agentes sociales y como titulares activos de sus propios derechos.

entonces la siguiente pregunta de investigación que busca responderse con la ejecución de este proyecto:

¿Cómo a partir de imágenes faciales puede ser estimada automáticamente la edad para controlar el acceso a contenidos permitidos en un sistema de control parental?

1.2. Justificación

Considerando los avances en la estimación automática de la edad a partir de imágenes faciales, algunos autores han planteado potenciales campos de aplicación, entre los que se encuentra la gestión electrónica de los clientes (ECRM²), donde por ejemplo los anunciantes pueden dirigir la publicidad de sus productos a clientes potenciales discriminando según la edad de los receptores o identificando automáticamente los grupos de edades de las personas que se muestran interesadas por los productos puestos en vitrina. Asimismo, estimar la edad a partir de los rasgos faciales con el fin de recoger datos demográficos y generar diversos informes relacionados con el perfil de los clientes de acuerdo a su edad e información útil para la inteligencia de negocios sin necesidad de obtener la identidad de los clientes ni violar su privacidad [3]. Por esto, han sido desarrolladas aplicaciones comerciales, como por ejemplo Releyeable [14], que analiza y captura automáticamente datos demográficos de los clientes para emplearlos en procesos de inteligencia de negocios, o FaceVACS-VideoScan [15] que mediante técnicas de reconocimiento facial realiza análisis demográfico, identificación facial, generación de estadísticas sobre métricas de audiencias, entre otras.

Otro campo de aplicación planteado es el control de seguridad y vigilancia, por ejemplo, un sistema con estimación automática de edad podría advertir sobre personas menores de edad que quieran adquirir un producto de venta restringida (tabaco, alcohol, etc.) en tiendas, bares o máquinas expendedoras, negar el acceso de ancianos a juegos mecánicos de alto riesgo [3] o restringir el acceso de los niños a sitios web para adultos o contenidos indebidos [4]. Sin embargo, a pesar del potencial e importancia de este campo de aplicación, no existe evidencia sobre la implementación y uso de un sistema de estimación automática de edad que satisfaga estas funciones.

Ahora bien, dada la importancia y la complejidad de estas situaciones adyacentes a los contenidos y los menores de edad, se encontró que existen soluciones software que han tratado de dar solución a este problema como por ejemplo los sistemas de control parental³.

² ECRM: *Electronic Customer Relationship Management*.

³ "Característica especialmente útil para padres y responsables educativos que desean impedir que niños o adolescentes puedan acceder a páginas Web inapropiadas" [133].

que son diseñados para restringir el acceso a contenidos inadecuados (principalmente para los niños), como es el caso de Net Nanny [16], la cual restringe y monitorea el uso del ordenador; y la aplicación para dispositivos móviles Qustodio Control Parental [17] que integra funciones similares. Al igual que éstas existen en el mercado varias aplicaciones que prestan servicios de este tipo, tanto para dispositivos móviles como para ordenadores [18] [19] [20]. No obstante, estas soluciones mencionadas utilizan mecanismos de seguridad, tal como las cuentas de usuario, donde los encargados de administrar el acceso a los menores, deben configurar manualmente las reglas que restringen el acceso al contenido.

En este contexto, es claro que en la actualidad existen soluciones software y servicios para el control parental que mitigan el problema de control de acceso a contenidos mediante funciones de monitoreo y bloqueo del material considerado inapropiado para el receptor. Sin embargo, dichas soluciones utilizan generalmente perfiles de usuario para configurar el acceso a aplicaciones y/o material específico, o utilizan funciones de bloqueo para su fin. Estas funciones deben configurarse de forma manual, por consiguiente, una solución que permita mediante técnicas de visión artificial determinar la edad del usuario y configurar de forma automática el control de acceso a contenidos inadecuados para su edad, podría ayudar a sustituir la configuración manual de las soluciones actuales, no solo para el control parental sino también para el control de acceso a lugares o artefactos que requieran una comprobación de edad para su uso. En este sentido, es necesario un mecanismo que combine las necesidades de las aplicaciones de control parental con la estimación automática de la edad a partir de imágenes fáciles para su configuración.

1.3. Objetivos

1.3.1. Objetivo general

Proponer una solución de software a nivel de prototipo funcional para el control parental a contenidos web mediante la estimación automática de la edad extraída de imágenes faciales.

1.3.2. Objetivos específicos

- Diseñar un mecanismo⁴ de estimación automática de edad que determine el grupo de edad al cual pertenece una persona a partir de una imagen de su rostro.
- Construir un prototipo funcional de sistema de control parental para el acceso a contenidos web desde un computador personal.
- Implementar e integrar el mecanismo de estimación automática de edad en el prototipo funcional de sistema de control parental.
- Evaluar la confiabilidad del prototipo de sistema de control parental mediante el esquema de evaluación LOPO⁵.

1.4. Aportes

A partir de la revisión de antecedentes que será presentada en el siguiente capítulo, y aunque existen propuestas tanto para el control de acceso a contenidos como para la estimación automática de la edad, esta propuesta se caracteriza a diferencia de las existentes porque combina ambas funcionalidades en un solo producto, es decir, existe una oportunidad con un alto potencial en la integración de estas funcionalidades, por lo tanto, en el presente trabajo se pueden identificar los siguientes aportes:

1.4.1. Investigativos

- Revisión de literatura de las principales técnicas de estimación automática de edad a partir de imágenes faciales.
- Mecanismo de Estimación Automática de Edad basado en características faciales.
- Integración de un mecanismo de estimación automática de edad en un prototipo de sistema de control parental.

1.4.2. De Innovación

- Contribución en el desarrollo de sistemas de control parental a partir del procesamiento de imágenes faciales que permitan la estimación automática de la edad.

⁴ Mecanismo: Un método o un sistema para lograr algo (Mechanism: A method or a system for achieving something). [<http://goo.gl/MnNjx9>]

⁵ *Leave-One-Person-Out*. Estrategia de evaluación popular sugerida en [5][138]

- Propuesta de una herramienta para proteger a menores de edad de contenidos web inapropiados en un computador personal.

1.5. Estructura del documento

En esta sección se describe la estructura de este documento, que contempla las diferentes etapas y teorías que permitieron el desarrollo del trabajo de investigación.

La primera sección es de carácter introductorio, donde son presentados el contexto y cuestión de investigación, junto a la justificación del problema. Adicionalmente, se exhiben los objetivos planteados que por ende conllevarán a la generación de aportes investigativos e innovadores, los cuales también son contemplados en esta sección.

En la *Sección 2 – Marco teórico y Antecedentes*, se introduce una base teórica y conceptual de los componentes de un sistema de estimación de edad. Incluyéndose también en esta sección una revisión de antecedentes, la cual presenta los trabajos considerados relevantes, además de los aplicativos que integran visión artificial y evidenciando algunas conclusiones parciales que orientarán este trabajo.

En la *Sección 3 – Mecanismo propuesto de la estimación de la edad*, se describen las técnicas y componentes utilizados que permitieron proponer un método de estimación de edad basado en los trabajos revisados y comparados con los antecedentes de la sección anterior. Se presentan los algoritmos que componen el método propuesto y el proceso a seguir para la implementación del mismo.

En la *Sección 4 – Evaluación del mecanismo propuesto*, se presenta un ciclo de evaluación del mecanismo propuesto de estimación de la edad mediante la técnica *Leave One Person Out* [5]. Aquí se muestran la planificación, diseño y resultados del desarrollo de la técnica para evaluar el método propuesto en términos de error medio absoluto (MAE).

En la *Sección 5 – Prototipo de control parental*. En esta sección se presentan los detalles de diseño del prototipo de control parental de acceso a direcciones URL, previamente almacenada en una lista de acceso organizada por categorías, además, se describen los componentes a integrarse en el sistema de control parental. Se incluye la planificación, diseño y resultados del desarrollo de software propuesto, usando para la descripción de las iteraciones, la metodología de desarrollo *Agile Unified Process* (AUP) [21].

Finalmente, en la *Sección 6 – Conclusiones y Trabajo futuro*, se presenta un resumen del trabajo de investigación realizado mediante la exposición de las conclusiones, producto del trabajo realizado y son sugeridas algunas propuestas como trabajos futuros de investigación.

2. MARCO TEÓRICO Y ANTECEDENTES

En esta sección se presenta la base teórica y conceptual en el cual se enmarco este trabajo investigativo y desarrollo del proyecto, adicionalmente son referenciados los trabajos revisados y relacionados con las temáticas de: control de acceso a contenidos, sistemas de control de acceso y seguridad, y técnicas de estimación de edad basadas en extracción de características de imágenes faciales, que son considerados antecedentes relevantes.

2.1. Marco teórico

2.1.1. Visión por computador

La visión por computador es el campo de investigación que estudia el análisis de imágenes a través de computadores para obtener una descripción de los objetos físicos que son captados por la cámara [22].

La visión por computador ha expandido su rango de aplicaciones en los últimos años debido a la aparición de dispositivos de captura de imágenes y de nuevos algoritmos. En la Tabla 1, son presentadas las principales aplicaciones de la visión por computador en algunos sectores de la industria y la ciencia.

Tabla 1. Áreas y ejemplos de aplicación de visión por computador.

Áreas	Ejemplo de aplicación
Industria Aeroespacial	Robots guiados mediante visión por computador y captura de imágenes telescópicas.
Industria Aeronáutica	Captura de Imágenes áreas en tiempo real y georreferenciadas.
Inspección de alimentos	Selección e inspección de calidad de alimentos según color, textura, formas, tamaños.
Automatización industrial	Inspección de embalajes, control de calidad y procesos de optimización de productos.
Industria automotriz	Detección de defectos en piezas y pintura de automóviles.
Identificación	Análisis de imágenes de huellas dactilares en laboratorios forenses y de investigación policial.
Deportes	Sistema de visión que permite mostrar jugadas controvertidas en partidos de futbol.

Medicina	Análisis de imágenes médicas con técnicas de recuperación de información para soporte a toma de decisiones en diagnóstico.
Multimedia	Aplicaciones multimedia interactivas, muestran proyecciones del cuerpo que permiten interactuar con multimedia (imágenes, sonidos).
Seguridad	Detección nocturna de peatones desde una distancia de 200m mediante sensores infrarrojos.
Transporte y tráfico	Sistema de visión para revisar el estado del pavimento.

En general, los sistemas de visión que se desarrollan en la industria intentan procesar imágenes sobre ambientes controlados (espacios bajo condiciones específicas de luminosidad), sin embargo, para áreas como el reconocimiento facial, se ha convertido en un reto detectar y procesar rostros en ambientes no controlados, de esta forma se han generado una gran cantidad de aplicaciones basadas en imágenes faciales.

Las imágenes faciales constituyen la entrada para el procesamiento digital en diferentes campos de aplicación, como el reconocimiento facial [23], en el que se determina la identidad de un individuo; la detección facial [3] en el que se intenta ubicar la dimensiones del rostro en la imagen; el análisis demográfico [24] intenta determinar la edad, raza, y el género al que pertenece un rostro; el análisis de video anónimo [25], intenta hacer una estimación en tiempo real de edad, raza, género y calcula métricas como la retención de la mirada sobre un punto, gestos faciales, parpadeos, entre otros, con el fin del análisis de clientes sobre sus preferencias; la cosmetología [26] utiliza imágenes faciales para realizar alteraciones por medio de filtros, para la mejoría estética del rostro; la biometría [3] facilita la medición y detección de puntos faciales; el control de seguridad y monitoreo de vigilancia [3] permite establecer si un individuo puede acceder a un espacio físico por medio de su identidad.

Por otra parte, el sexo y la edad estimado mediante una imagen facial ha sido incorporado en áreas diferentes al reconocimiento de patrones, tal como ocurre en la inteligencia de negocios, en donde se recolecta información sobre los gustos de clientes que varían de acuerdo a la edad y sexo, de manera que esta información se puede utilizar en beneficio de las empresas.

En este contexto, la estimación de la edad en particular, tiene potenciales aplicaciones de desarrollo [27], entre ellas: restringir el acceso mediante puertas provistas de sensores que alerten sobre el ingreso a establecimientos prohibidos por edad; restricción de control parental con información de edad (que es el tema principal de este trabajo); vigilar acceso a áreas mediante sistemas de cerrados de vigilancia; alertar sobre transacciones sospechosas en cajeros electrónicos; organizar fotografías de una persona para realizar búsquedas con información específica; filtrar imágenes para motores de búsqueda, las

imágenes resultantes de una consulta podrían ser filtradas por edades facilitando y acelerando el proceso de búsqueda; identificación humana transversal a la edad, la estimación de la edad puede ayudar mejorar el reconocimiento facial.

2.1.2.Representación de una imagen

Las imágenes capturadas por cámaras digitales son representadas por una función bidimensional $f(x, y)$, en donde el valor de la función representa el grado de intensidad de luz en el espacio de coordenadas (x, y) . Las imágenes se representan en 3 dimensiones con dominio y rango discreto. Una imagen es almacenada en una matriz de $N \times M$ elementos tal como se presenta en la ecuación (1). La representación digital en la memoria de un computador se comporta como una matriz, donde el valor de los elementos es su nivel de gris (pixel), la fila y la columna su posición en el espacio [22].

$$f(x, y) = \begin{bmatrix} f(0, 0) & f(1, 0) & \dots & f(N-2, 0) & f(N-1, 0) \\ f(0, 1) & f(1, 1) & \dots & f(N-2, 1) & f(N-1, 1) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ f(0, M-2) & f(1, M-2) & \dots & f(N-2, M-2) & f(N-1, M-2) \\ f(0, M-1) & f(1, M-1) & \dots & f(N-2, M-1) & f(N-1, M-1) \end{bmatrix} \quad (1)$$

Algunas definiciones [22] sobre el color de la imagen son:

- El **brillo** es la sensación que indica si un área está más o menos iluminada.
- El **tono** es la sensación que indica si un área parece similar al rojo, amarillo, verde o azul o a una proporción de dos de ellos.
- La **coloración** es la disposición y grado de intensidad de los diversos colores.
- La **luminosidad** es el brillo de una zona respecto a otra blanca en la imagen.
- La **croma** es la coloración de un área respecto al brillo de un blanco de referencia.
- La **saturación** es la relación entre la coloración y brillo.

Las tres dimensiones de la imagen determinan el nivel de gris que tienen cada componente, y la suma de estas componentes corresponden al nivel de luminosidad de la imagen. Este modelo de color se llama RGB (*Red, Green, Blue*), y representa la cantidad determinada de cada dimensión de color que forman un color. Cada dimensión del modelo RGB varía entre 0 y 255 (**escala de grises**).

Los tres elementos que componen un pixel en el modelo RGB se pueden representar en una escala de grises, con el objetivo de facilitar operaciones sobre matrices y disminuir el costo de procesar cada componente de color. La fórmula que permite hacer esta transformación se describe en la ecuación (2).

$$g = c_1R + c_2G + c_3B \quad (2)$$

En donde R, G, B , es la intensidad de cada componente de color, y c_1, c_2, c_3 , representa el porcentaje de cada componente, de modo que $c_1 + c_2 + c_3 = 1$. Al reemplazar los valores por: $c_1 = 0.299, c_2 = 0.587, c_3 = 0.114$, se obtiene una imagen en escala de grises [28], en donde los valores representan el 0 como ausencia de intensidad de color (negro), y 255 como el máximo de intensidad (blanco). La imagen en escala de grises se puede representar en forma de gráfico, calculando las frecuencias de aparición de los valores de gris. Este gráfico se denomina histograma de la imagen y permite entre otras cosas segmentar objetos en algunos casos. En la Figura 1 se muestra el histograma de una imagen obtenida de la base de datos de imágenes faciales FG-NET [29].

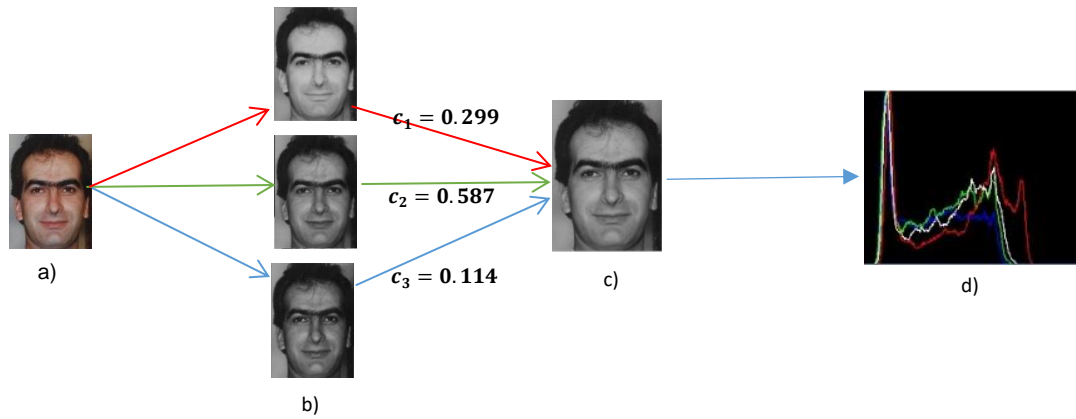


Figura 1. Histograma de una imagen. a) Imagen original $f(x, y)$, b) componentes de color de la imagen original RGB, rojo, verde y azul de arriba abajo. c) Imagen en escala de grises $g(x, y)$. d) Histograma de la imagen, la línea en blanco representa el histograma para la imagen en escala de grises.

Con la información por píxeles obtenida del histograma, es posible obtener algunos píxeles que están por encima o debajo de cierto nivel de intensidad. Ese nivel de intensidad se define como umbral, en la ecuación (3), describe el proceso de obtención de una imagen umbralizada a partir de una imagen en gris.

$$F'(x, y) = \begin{cases} 0, & \text{gris}(x, y) < U \\ 255, & \text{gris}(x, y) \geq U \end{cases} \quad (3)$$

Donde $gris, F'$, corresponden a las imágenes en escala de grises y la umbralización respectivamente, el umbral U es el nivel de intensidad que permite la binarización (imagen con dos intensidades 0 y 255).

Dada una imagen en escala de grises puede mejorarse la apariencia de la imagen original mediante transformaciones que cambian el contraste (diferencia entre los píxeles claros y los oscuros) de la imagen. La modificación del contraste se realiza de dos formas, la primera se trabaja sobre la transformada de Fourier de la imagen, y en la segunda se modifican los

pixeles directamente. La forma directa modifica los pixeles independientemente del resto de pixeles de la imagen.

Con respecto al **ruido** en una imagen, este se puede definir como aquellos pixeles generalmente aislados con niveles de gris diferentes al de sus vecinos.

Otro término que se hace necesario definir, es el **filtro** (también conocido como máscara, *kernel*, plantilla, ventana o matriz de convolución), el cual es una matriz de coeficientes que es operada con la imagen en forma de ventana móvil centrada en cada pixel de la imagen. Esta operación matemáticamente se conoce como una **convolución** de la imagen con la matriz. En la ecuación (4), $I'(i, j)$ es la imagen filtrada, m, n , son las dimensiones de la imagen, K es la matriz de convolución.

$$I'(i, j) = \sum_{m,n=-1}^{m,n=1} I(i + m, j + n) \cdot K(m + 1, n + 1) \quad (4)$$

Los filtros que eliminan el ruido y otros pequeños detalles en la imagen que no son de interés se denominan de paso bajo. Mientras que los filtros de paso alto realzan detalles de la imagen. En el caso de la detección de los bordes primero se debería pasar por un filtro de paso bajo para realzar la diferencia de intensidades en el borde con un filtro de paso alto.

2.1.2.1. Transformaciones de imágenes basadas en su vecindad

La **transformación** de una imagen se logra al operar cada pixel de la imagen basado en su vecindad para determinar un valor de la nueva imagen. Las transformaciones son necesarias cuando las imágenes necesitan pasar por nuevas etapas de procesamiento para obtener mejoras en su contraste, brillo, etc. En la Figura 2, se muestra la ampliación de una imagen y una pequeña región donde se resalta un pixel y su vecindario.

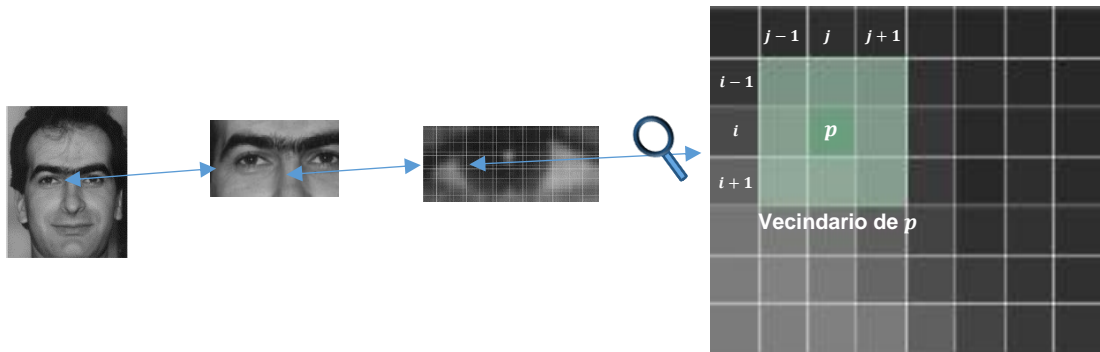


Figura 2. Ampliación de una imagen y vecindario de un pixel.

Cada pixel de la imagen presenta un total de ocho vecinos (Figura 2), por ejemplo en la posición de coordenadas (i, j) , los vecinos del pixel p son:

$i - 1, j - 1$	$i, j - 1$	$i + 1, j - 1$
$i - 1, j$	$p(j, i)$	$i + 1, j$
$i - 1, j + 1$	$i, j + 1$	$i + 1, j + 1$

Figura 3. Coordenadas de los vecinos de un pixel p .

Las transformaciones sobre la imagen basan sus cálculos sobre el vecindario del pixel p [22]. De manera que se definen los vecinos del pixel p (ubicado en la posición i, j), en forma lineal en la ecuación (5).

$$Vecinos_{p(i,j)} = \{p(i - 1, j - 1), p(i - 1, j), p(i - 1, j + 1), p(i, j + 1), p(i + 1, j + 1), p(i + 1, j), p(i + 1, j - 1), p(i, j - 1)\} \quad (5)$$

El cálculo del valor de un pixel de la imagen transformada depende de la influencia de los niveles de intensidad de los ocho vecinos, según la ecuación de una convolución (4), en donde la matriz de convolución determina en qué medida contribuyen los vecinos a la modificación del pixel central.

2.1.3.Descriptores de texturas

El concepto textura está relacionada con la suavidad, rugosidad o regularidad de una imagen, sin embargo, no hay una definición formal de textura. Algunos autores han tratado de definir la textura desde una perspectiva concreta de su naturaleza [30], como:

- Región organizada que puede ser descompuesta en primitivas teniendo unas distribuciones espaciales concretas.
- Región de la imagen bidimensional, aleatoria y posiblemente periódica.
- Distribución espacial de color o intensidad en una imagen o en una región de la misma.

Así, las texturas se pueden clasificar en diferentes categorías dependiendo del problema que se desee resolver:

- Microtexturas o macrotexturas.
- Irregulares o regulares.
- Repetitivas o no repetitivas.
- Con alto contraste y no direccionales frente a bajo contraste y direccionales.

- Granulada y de baja complejidad frente a no granuladas y de alta complejidad.

Ahora bien, un descriptor es una representación de la distribución de píxeles en una región de una imagen, que permite obtener valores que representen la forma en la cual aparecen las diferentes intensidades de color de una imagen, en el caso de las texturas los descriptores se han clasificado en las siguientes categorías [22]:

- Estructurales:** Representan la textura por medio de una jerarquía y primitivas bien definidas.
- Estadísticos:** Representan propiedades basadas en la distribución y la relación entre los niveles de gris de la imagen.
- Basados en modelos:** Representan la textura mediante fractales o modelos aleatorios.
- Basados en transformaciones:** Representa la imagen en un espacio con un sistema de coordenadas relacionado a las características como la frecuencia o el tamaño de la textura.

2.1.4. Sistema de reconocimiento de patrones

Un patrón es un conjunto de objetos, fenómenos o conceptos donde los elementos del conjunto son similares entre sí, en ciertas formas o aspectos [31]. Entre algunos ejemplos de patrones se tienen: las huellas dactilares, formas de árboles, frecuencias en señales digitales, rostros, tipografías, códigos de barras, etc. Los patrones se pueden agrupar en un conjunto con propiedades comunes entre sí, esto es conocido como categoría de patrones.

Según Bishop [32], el reconocimiento de patrones es el campo que estudia el descubrimiento automático de regularidades en los datos mediante el uso de algoritmos informáticos y el uso de las regularidades para tomar acciones tales como clasificar datos en diferentes categorías de patrones. En la Figura 4 se muestra un diagrama de bloques de este modelo de reconocimiento de patrones.

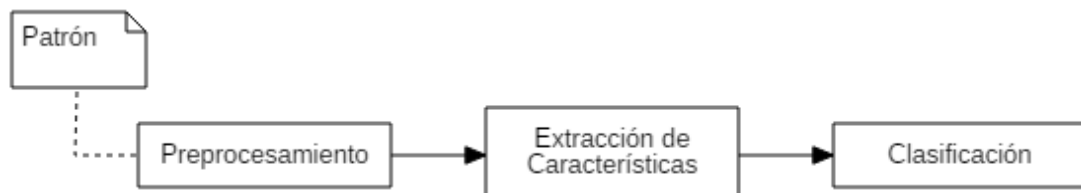


Figura 4. Diagrama de bloques de un modelo estadístico de reconocimiento de patrones. Tomado de [31]

Un sistema de reconocimiento de patrones involucra en esencia 3 pasos, tal como se muestra en la Figura 4. El preprocesamiento es la etapa donde se obtienen los datos del

ambiente (por ejemplo, imágenes obtenidas por cámara web), que servirán de entrada al sistema de reconocimiento de patrones. A estos datos de entrada se les elimina el ruido y son extraídos patrones de interés, de tal forma que los datos sean entendibles por el sistema. En la siguiente etapa se extraen las características relevantes de los datos procesados, a partir de la cual se obtiene una entidad del objeto que va a ser reconocido o clasificado. Finalmente, se realiza la etapa de la clasificación, en donde se decide la clase a la que pertenece una muestra según el vector de características.

Los modelos de reconocimiento de patrones se han categorizado para Sharma e Kahun [31]:

- **Modelo estadístico:** Modelo basado en estadística y probabilidad. El patrón es descrito en términos de conjuntos de características. Éstos conjuntos son seleccionados de tal manera que no se solapan entre sí, y la efectividad de la selección se determina escogiendo la mejor forma de separar los patrones de diferentes clases entre sí. Las características son seleccionadas después de analizar los patrones de entrenamiento. Los sistemas entrenan desde los patrones de entrenamiento y se adapta para reconocer o clasificar los patrones de evaluación desconocidos.
- **Modelo estructural:** Están basados en la relación entre características. Los patrones son representados por estructuras que pueden tomar en cuenta relaciones más complejas entre características. También son llamados modelos sintácticos. Aquí los patrones a ser reconocidos son llamados “**primitivas**”, y los patrones complejos son representados por la relación formada entre esas primitivas y las reglas gramaticales asociadas con esa relación. Aquí los patrones son las sentencias pertenecientes a un lenguaje, las primitivas son el alfabeto del lenguaje, y mediante el uso de esas primitivas, las sentencias son generadas de acuerdo a la gramática. Este paradigma es usado en situaciones donde se tiene una estructura definida y además, pueden ser capturadas en forma de un conjunto de reglas.
- **Modelo de emparejamiento de plantillas:** Este modelo es usado para determinar la similitud entre dos imágenes, píxeles o identificar curvas y figuras. Una imagen de entrada es comparada con un patrón conocido (plantilla), mediante la maximización de la correlación cruzada espacial o minimizando una distancia. Cada pixel es comparado teniendo en cuenta todas las posibles posiciones, rotaciones y cambios en la escala de la imagen de entrada respecto a la plantilla.
- **Modelo de redes neuronales:** Este modelo funciona bajo el paradigma de una red neuronal artificial, considerada como un proceso entrenable auto-adaptativo que es capaz de aprender y resolver problemas complejos basados en conocimiento disponible. Éste modelo intenta usar algunos principios organizacionales tales como el aprendizaje, la generalización, la adaptabilidad, la tolerancia a fallos, representación distribuida y computación en la red de grafos dirigidos con pesos en el cual las neuronas forman los nodos y los arcos (con pesos) son las conexiones entre las neuronas de salida y, de entrada. Los pesos aplicados a los resultados de

las conexiones desde el aprendizaje indican la importancia de la contribución de la neurona precedente en la información que va ser pasada a la siguiente neuronal.

2.1.5. Máquinas de soporte vectorial

Una máquina de soporte vectorial (SVM⁶) es un algoritmo de clasificación que construye un hiperplano n-dimensional para separar los datos óptimamente en dos categorías [31]. Una SVM es aplicada para resolver tareas tanto de clasificación como de regresión. La solución estándar que realiza este algoritmo permite una clasificación lineal binaria no probabilística (dos clases apartadas por un hiperplano o separador lineal), el cual toma un conjunto de datos de entrada y predice las clases posibles, por cada entrada, entre las dos posibles clases. Otro tipo de problemas que resuelve SVM son los problemas cuasi-separables (no linealmente separables, pero separables en un espacio transformado, mediante funciones *kernel*).

Una SVM para clasificación binaria de ejemplos linealmente separables, busca un hiperplano denominado óptimo, si su margen es de tamaño máximo [33]. En la gráfica izquierda de la Figura 5, se muestra un conjunto de punto etiquetados para entrenamiento, pertenecientes a dos clases que se separan por uno de los infinitos planos que esta entre el espacio donde no hay elementos entre los puntos, sin embargo, el SVM busca minimizar el riesgo estructural [33], por tanto el hiperplano óptimo maximizado calculado por SVM es el hiperplano de margen con tamaño máximo y se muestra en la parte derecha de la figura.

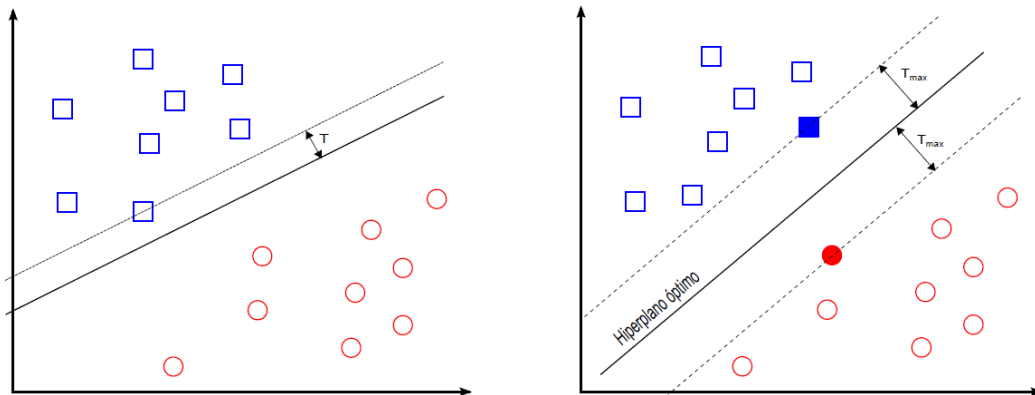


Figura 5. Hiperplano óptimo entre dos clases (azul y rojo) de muestras. Imagen tomada de [34].

⁶ Support Vector Machine

Las clases dibujadas pertenecen a un conjunto separable de ejemplos $S = \{(x_1, y_1), \dots, (x_n, y_n)\}$, donde $x_i \in \mathbb{R}^d, y_i \in \{-1, +1\}$, (y_i representan las etiquetas de las muestras de entrenamiento), se puede definir el hiperplano (6) de margen con tamaño máximo, como una función lineal que es capaz de separar el conjunto S , sin error.

$$D(x) = \langle w, x \rangle + b \quad (6)$$

Aquí w y b , son coeficientes reales. Se debe cumplir con restricción descrita en la ecuación (7).

$$y_i D(x_i) \geq 0, \quad i = 1, \dots, n \quad (7)$$

Sin embargo, la restricción anterior no es suficiente para cumplir la separación óptima, por lo que su margen τ (mínima distancia equidistante entre el hiperplano y los ejemplos más cercanos de cualquiera de las dos clases) debe ser máximo (gráfica izquierda de la Figura 5). Por tanto, se debe cumplir que la distancia entre el hiperplano de separación y los ejemplos del conjunto S sean máximas y el vector w sea perpendicular (8).

$$y_i D(x_i) \geq \tau \|w\| \quad (8)$$

Reciben el nombre de **vectores soporte** los ejemplos que están situados a ambos lados del hiperplano óptimo y que definen el margen o, lo que es lo mismo, aquellos para los que la norma de w es mínima, es decir, un margen máximo. Entonces, se define formalmente en la ecuación (9), el problema de encontrar el valor de w y b , que minimiza a $f(w) = \|w\|$. Éste puede ser resuelto usando multiplicadores de Lagrange para obtener el vector y el margen del hiperplano óptimo.

$$\min f(w) = \frac{1}{2} \|w\|^2 = \langle w, w \rangle, \text{ siempre y cuando } y_i D(x_i) \geq 1, \forall i \quad (9)$$

El método SVM es utilizado de una forma semejante como un método eficiente de búsqueda de un hiperplano óptimo para separar datos no lineales. En la actualidad, el SVM tradicional también es usado en clasificación de múltiples clases donde los datos a ser clasificados pueden pertenecer a una o más clases entre un número de clases mediante funciones *kernel* y transformaciones de espacios [34].

Entre las ventajas de usar SVM pueden mencionarse [35]:

- Facilidad de entrenamiento: Dado un conjunto de muestras de entrenamiento donde cada una de ellas es marcada como perteneciente a una de las dos categorías, el algoritmo de entrenamiento construye un modelo que asigna nuevas muestras en una categoría o la otra.
- No hay un óptimo local, como sucede en los algoritmos de aprendizaje no supervisado (redes neuronales).

- Para clasificaciones múltiples en espacios transformados, se escala bien en dimensiones altas.
- El compromiso entre la complejidad del clasificador y el error puede ser controlado explícitamente.
- Una gran cantidad de datos estructurados pueden ser usados como entrada a SVM, en lugar de vectores de características.

La selección de un *kernel* adecuado es una gran desventaja que demuestra éste método en el caso de espacios transformados para clasificación de datos no linealmente separables, es decir, una metodología eficiente para sintonizar los parámetros de inicialización del algoritmo[36].

2.1.6. Estimación automática de edad

La estimación automática de edad es la tarea de predecir un valor estimado de la edad de una persona basado en su imagen facial. La literatura evidencia que en los últimos años se han generado técnicas de estimación utilizando nuevos algoritmos y variantes de los ya existentes, mejorando en su precisión y rendimiento. Éstas técnicas generalmente vienen estructuradas en dos módulos principales [3], el primero que comprende la representación de la cara, en el cual los rasgos faciales relacionados con la edad humana se extraen de los rostros para obtener una representación compacta desde una imagen digital; y el segundo es la aplicación de técnicas para la estimación de la edad, en el cual se toma como base las características extraídas para realizar dicha estimación. En la figura 6, es presentado mediante un diagrama de bloques este proceso de estimación de edad.

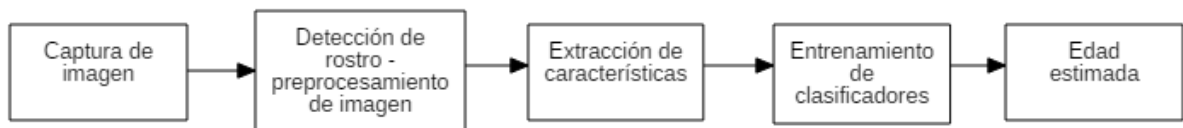


Figura 6. Proceso de un sistema de estimación de edad.

En el proceso de estimación de edad descrito en la Figura 6, inicialmente se capturan las imágenes de entrada por medio de una cámara web o un video; cada imagen se mejora en contraste, se suaviza, y se realizan otras operaciones propias del procesamiento digital; esta imagen pre-procesada entra al detector de rostros y cada detección positiva es plasmada en una nueva imagen. Por medio de un algoritmo de extracción de características que recibe cada rostro, se extrae un vector de características, el cual después es evaluado por un algoritmo clasificador de los grupos de edad. Posteriormente, la etapa de clasificación, se inicia con el entrenamiento del clasificador, para luego en la etapa de prueba clasificar las imágenes en cada uno de los grupos de edad. Los grupos de edad son usados para afinar la estimación de la edad, es decir, que la determinación de estos grupos de edad influye en la precisión de la estimación (por ejemplo, entre más amplio sea el rango

del grupo de edad, se obtienen resultados cercanos a los esperados). Cada una de las etapas del proceso de estimación descritas, está conformada por un conjunto de algoritmos que determinan el éxito de la estimación.

2.2. Antecedentes

Las principales técnicas de estimación de la edad se dividen en cuatro ramas, en primer lugar, la clasificación donde se considera cada etiqueta de edad como una clase y el patrón de envejecimiento puede ser aprendido por algoritmos de clasificación o clasificadores. Muchos de esos algoritmos han sido desarrollados basados en el aprendizaje de máquina, como por ejemplo las redes neuronales artificiales (ANN⁷) [37] y las máquinas de soporte vectorial (SVM⁸) [38]. En segundo lugar, la edad puede ser considerada como un problema de regresión en el que se pueden emplear técnicas como el regresor de soporte vectorial (SVR⁹) y el análisis de correlación canónica (CCA¹⁰) [39]. En tercer lugar, encontramos los enfoques híbridos que fueron diseñados para combinar las ventajas de los métodos de clasificación y regresión como lo podemos apreciar en [4] y [5]. En estos trabajos se propone el regresor robusto ajustado localmente (LARR¹¹). Por último, el rango o “*ranking*” introducido en [40], considera que el proceso de envejecimiento facial es un procedimiento no estacionario. En este sentido, se define que la diferencia entre edades como por ejemplo 52 y 51 no es igual a la diferencia existente entre 22 y 21. Por lo tanto, la información que se usa es el orden relativo entre las etiquetas de edad, en lugar de sus valores exactos.

2.2.1. Modelos de representación facial

Una de las primeras investigaciones dirigidas hacia la estimación de la edad la realizó Kwon y Lobo [41], clasificando las imágenes faciales en tres categorías: bebés, adultos jóvenes y adultos mayores, basándose principalmente en el modelo matemático del crecimiento craneo facial [3]. Liu et al. propusieron un clasificador que consta de dos etapas [42], cada una de éstas compuesta de clasificadores y extractores de características, obteniendo como resultado una precisión promedio de 73.3% en los tres grupos de clasificación (bebé, niño y adulto) con respecto a la propuesta de Kwon y Lobo [41], y reduciendo en 56% la determinación del grupo de edad. Dicho enfoque de dos etapas [43], amplía el número de grupos en cinco, mediante el uso de seis características geométricas basados en radios

⁷ *Artificial Neuronal Networks*

⁸ *Support Vector Machine*

⁹ *Support Vector Regression*

¹⁰ *Canonical Correlation Analysis*

¹¹ *Local Adjusted Robust Regressor*

sobre la imagen facial, y mediante los clasificadores LDC, NNC y SVC, mejoran los resultados en cuanto a precisión, rapidez y eficiencia con respecto a los anteriormente expuestos en un 25.22%.

2.2.1.1. Modelos de Apariencia Activa

Los modelos de apariencia activa (AAM) [44], permiten extraer características globales y locales dado un conjunto de entrenamiento sobre el que se realiza análisis estadístico de la figura y apariencia sobre las escalas de gris de la imagen. Liu et al. en lugar de dividir la información de figura y textura, las combina en un solo vector, seguido a esto les reduce su dimensionalidad mediante PCA y finalmente pasan las imágenes sobre el clasificador de 5 grupos [45]. Sin embargo, más adelante se plantean una debilidad aparente de AAM [46], ya que se considera que el método es bueno principalmente en la extracción de características globales resultante en una buena cantidad de información de la apariencia y la figura del rostro, pero no se obtiene información sobre arrugas y envejecimiento de la piel; de manera que se propone un método para extraer información a partir de texturas por medio de extracción de características globales mediante el AAM, y mediante filtros de Gabor para las características locales, así, logran conseguir un resultado MAE de 4.6577. Éste modelo es comúnmente utilizado, debido a que se pueden calcular los puntos de contorno de la cara y con ellos extraer el rostro del resto de la imagen, como se puede apreciar en los trabajos de Li et al. [47] y Gao [48]¹².

2.2.1.2. Subespacio del patrón de envejecimiento (AGES¹³)

El método AGES [49] utiliza secuencias del rostro a través de diferentes etapas de la vida de modo que representa el proceso de envejecimiento (patrón de envejecimiento). Al momento de formar el patrón de envejecimiento puede faltar cierta imagen en alguna de las etapas en orden temporal, AGES puede sintetizarla aplicando algoritmos de aprendizaje iterativo como el algoritmo EM¹⁴. AGES tiene la particularidad de requerir de una entrada de secuencia de imágenes faciales con su identificación, sin embargo, las bases de datos actuales pueden no satisfacer ese requerimiento por lo que resulta en carencia de información de texturas.

¹² [48] extrae las características utilizando para su implementación AMTools. <https://goo.gl/j3z8xt>.

¹³ *Aging Pattern Subspace*

¹⁴ *Expectation maximization algorithm.*

2.2.1.3. Age Manifold

El método *Age Manifold* [5], aprende un patrón de envejecimiento común sobre un gran número de imágenes de entrenamiento de múltiples individuos en diversas edades. Los datos de la imagen facial original son asignados a un sub espacio de baja dimensión (colector), y ajustados con una función de regresión robusta (LARR). Para la evaluación de una imagen, se debe detectar y normalizar la imagen, lograr que sea proyectada sobre el colector aprendido, mediante la función de regresión se logra predecir la edad.

2.2.1.4. Los filtros de Gabor y el operador Patrón Binario Local

Los filtros de Gabor son capaces de determinar la dirección dominante de las arrugas, y extraen efectivamente características de arrugas por región [46]. Los métodos que combinan filtros de Gabor con descriptores locales han demostrado que obtienen buenos resultados de estimación [46] [50].

El método patrón binario local (LBP¹⁵) permite extraer texturas detalladas de piel (características de piel). [51]. El operador LBP inicialmente es ampliamente usado como un descriptor de características debido a su robustez frente a variaciones de luz, pose e iluminación [52].

Los filtros de Gabor y el LBP son representaciones ricas en información y computacionalmente eficientes, además su carácter complementario los convierte en buenos candidatos para fusionarlos y lograr un mejor resultado.

2.2.1.5. Características Biológicamente Inspiradas

El método características biológicamente inspiradas (BIF¹⁶) [53] es empleado como extractor de características de envejecimiento facial locales y globales, éste método se basa en el modelo HMAX que muestra buen rendimiento en detección de objetos. Éste modelo puede capturar datos como arrugas en la frente y flacidez e inclinación de los músculos faciales (características locales), y figura cráneo facial y radios de características faciales (características globales). El método BIF logra una representación con características de envejecimiento de dimensionalidad 8940, por lo que se debe pasar un reducto de dimensionalidad (LDA) para que baje a 150, esta reducción es esencial para la estimación puesto que de esta manera el trabajo presentado por Chen et al. [53] junto a CS-OHRank

¹⁵ *Local Binary Pattern*

¹⁶ *Bio-Inspired Features*

logran uno de los mejores métodos de estimación para características biológicamente inspiradas.

2.2.1.6. Los métodos basados en redes neuronales convolucionales (CNN)

A diferencia del método BIF [53], los métodos basados en redes neuronales convolucionales tienen una estructura más profunda y con la capacidad de aprender sus parámetros. Este método de estimación a comparación con los anteriores no cuenta con el paso intermedio entre la extracción de características y la regresión, de esta manera se optimizan para minimizar el error en la estimación. Yi et al. [54] proponen un enfoque multitarea que permite estimar además de la edad, el género y la clasificación étnica.

Qiu et al. [55] utilizan CNN en las capas bajas de la arquitectura de su sistema propuesto. Este método ha sido probado en MORPH-II y han obtenido un MAE de 4.43, superando a los métodos con los que se han comparado.

Algunos autores [56] [57], afirman que la mayoría de los trabajos presentados hasta el momento han sido probados efectivamente con técnicas (*benchmarks*) de alcance restringido (imágenes faciales en condiciones controladas) para la estimación de edad. Debido a lo anterior dichos autores proponen usar técnicas que abarquen un espectro mayor de condiciones menos controladas, con la ayuda de imágenes extraídas de redes sociales, debido a que se ha vuelto difícil de interpretar resultados casi perfectos de MAE con diferencias tan mínimas.

2.2.1.7. Métodos basados en rangos

Otras formas de clasificación diferentes a las basadas en clases y regresión han obtenido buenos resultados en investigación recientemente, tal es el caso de OHRank [58], en donde la idea principal es el orden relativo de la edad humana. De modo que se divide la secuencia de envejecimiento completa en dos partes usando hiperplanos para determinar la edad a la que pertenece una muestra (imagen). Éste trabajo logra un MAE de 4.48. El método OHRank es usado por Li et al. [47] quienes plantean una refinación de la estimación mediante asignación de pesos y la implementación de la clasificación basada en representación dispersa dividiendo las edades en grupos pequeños, de manera que en una etapa posterior se pueda estimar la edad con una mejor precisión, en comparación con la propuesta de Chang et al. [58], se mejora el MAE con un resultado de 4.32. Más tarde el método OHRank es implementado y mejorado por Chang y Chen [59] superando los resultados de clasificadores convencionales de regresión y ranking.

El método descrito por Weng et al. [60] resalta que diferentes características pueden reflejar variados efectos de envejecimiento, motivados por esto, el método MFOR extrae características únicas desde vistas diferentes para representar la información de la edad, así divide el problema en K-1 subproblemas convirtiéndolo en un problema de rangos. Los

resultados del método en la base de datos MORPH-II es de un MAE de 4.20, superando a otros métodos basados en rangos como OHRank.

2.2.1.8. Procesos Bayesianos

Los procesos bayesianos [61] interpretan por medio de la probabilidad diferencias en la intensidad de la imagen. Éstos métodos que intentan resolver el problema de la cantidad de datos de entrenamiento por cada persona que requieren y la falta de precisión para asignar un estimador a una imagen facial [62].

El trabajo presentado por Gao et al. [63] consta de dos regresores de procesos bayesianos, el proceso gaussiano y *t-process*, aquí los parámetros en los modelos pueden ser aprendidos desde los datos automáticamente usando métodos de gradiente eficientes. *El t-process* es complementario al proceso Gaussiano ya que adiciona control de robustez al problema que afecta al proceso gaussiano por los valores atípicos.

2.2.1.9. Distribución de etiquetas

Dado que los rostros de edades cercanas son parecidos. En lugar de considerar cada imagen facial como una muestra con una única edad (etiqueta), se asocia una distribución de etiquetas que cubre un número de clases de etiquetas representando el grado que cada etiqueta describe la muestra [64]. Este método (IIS-LLD) aprende automáticamente las distribuciones de etiquetas a diferentes edades e intenta solucionar el problema mencionado con anterioridad a los parecidos a edades cercanas. Los autores demuestran que se obtienen más ventajas con distribuciones que sobre los métodos basados en una sola etiqueta (edad) como el método AGES.

Yin y Geng [65] intentan resolver el modelo de entropía máximo [64] por medio de redes neuronales que puedan aprender automáticamente la probabilidad condicional de la edad dada la imagen facial.

2.2.1.10. Enfoque de Fusión de características

En esencia una fusión de características consiste en mezclar dos o más métodos de extracción con buena aceptación para formar uno más preciso. Éstos generalmente logran buen rendimiento, robustez y mejoras significativas del sistema de estimación de edad (comparados con un sistema de estimación de un solo extractor). Por ejemplo: Chen et al. [51] fusionan las características locales y globales mediante Gabor más LBP o Gabor más AAM [50]. Bajo el mismo enfoque Liu et al. [66] extraen características locales y globales por medio de BIF, HOG y LBP, a partir de éstas se forman seis esquemas de fusión de características con el fin de mejorar el rendimiento, se destaca un MAE de 2.97 y 2.81 (para MORPH-II y FG-NET respectivamente) [65]. Chen et al. [51] evalúan el rendimiento de 39

combinaciones diferentes de cuatro métodos de normalización de características, se presentan dos métodos de fusión, con representaciones de AAM, Gabor y LBP [50]. Inspirados por el trabajo de Chen et al. [52] fusionan características globales (AAM) y características locales (Gabor) como la representación de los rostros [51].

En el enfoque que presentan Feng et al. [67] se extraen los puntos característicos de atributos de textura y figura (AAM y LBP), y se fusionan mediante el análisis de correlación canónico (CCA). De esta manera logran mejor precisión en la estimación y menos tiempo de computación superando a métodos como el BIF.

La fusión de características locales y globales propuesta por Pontes et al. [68], son combinadas para obtener un nuevo vector de características híbrido con información de las arrugas, textura, manchas faciales (LBP, Gabor y LQP), e información proporcionada por la figura y apariencia (AAM). En la etapa de estimación se clasifica en un grupo de edad determinado basado en SVM y con base en tal resultado se define una edad específica mediante SVR, de manera que se pueda mejorar para obtener un mejor resultado.

2.2.1.11. Resumen de resultados de precisión

A continuación, se presentan de forma consolidada mediante las Tabla 2, 3 y 4, los trabajos de investigación con sus respectivos resultados en precisión y el error medio absoluto que representa la medida de comparación entre métodos. Los resultados presentados corresponden a trabajos de investigación elaborados en los años siguientes a la publicación de Fu et al. [3], por lo que es un complemento de dicho trabajo. Las bases de datos que son utilizadas en la fase de entrenamiento y evaluación son la FG-Net y la MORPH (álbum 2). La columna "Id", representa la referencia de cada artículo, en la columna "Extracción y Selección" se listan los métodos que se utilizaron en la etapa previa a la clasificación, en la que se extraen las características y ocurren procesos de reducción de dimensionalidad, la columna de "clasificación" presenta los métodos utilizados para obtener el resultado de la estimación bien sea como clasificación o como una etapa de regresión. Finalmente, las columnas de "MAE" en unidades de edad, y "Grupos de Clasificación" representa las clases que se utilizaron para la etapa de clasificación y se marca con *EE* en caso que se presente un caso de regresión.

A seguir, en la Tabla 2, se presentan los trabajos que utilizaron la base de datos FG-NET para entrenamiento y evaluación. La base de datos FG-NET está compuesta por imágenes con variadas condiciones de luz, y con objetos como gafas, barbas que obstaculizan la visión completa del rostro, además de algunas imágenes con baja calidad de resolución, por tanto, esta base de datos es utilizada para simular un ambiente no controlado.

Tabla 2. Resumen y comparación de técnicas de estimación de edad sobre la base de datos FG-NET.

Id	Extracción y Selección	Clasificación	MAE	Grupos de Clasificación
[69]	Pseudo Zernike Moment (PZM)	k-NN	8.75	0-10, 11-20, 21-30, 31-40, 41-50
		SVM	4.56	
		Multi-Layer Perceptron (MLP)	1.86	
[70]	Gabor Filters, LBP, GLCM statics, PCA y MvDA (reducción dimensionalidad)	SVM	5.41	0-2, 3-7, 8-12, 13-19, 20-36, 37-65, >66
[67]	AAM, LBP	CCA, Extreme Learning Machine (ELM) con un kernel Gausiano (RBF)	5.55	EE
[71]	completed local binary pattern (CLBP)	SVR	5.09	EE
[61]	AAM	Gaussian Process, t Process	5.23	EE
[72]	Affine-invariant shape representation, AAM	SVM	5.89	EE
		Ridge Regression	7.57	
		RVM	6.69	
[73]	multi-block local binary pattern (MB-LBP)	implementación propia	4.94	EE
[53]	BIF, LDA (reducción dimensionalidad)	CS-OHRank	4.72	EE
[46]	Gabor filters, LBP, AAM	SVM, AdaBoost y LDA	4.6577	EE
[74]	Lie Algebrized Gaussians (LAG)	SVR, WCCN basado en SVM	4.54	EE
[68]	AAM, LBP, GABOR WAVELETS, LPQ	SVM, SVR	4.5	EE
[48]	AAM	Clustered Multi-task Support Vector Regression Machine CMTSVR	4.37	EE
[75]	CS-LBFL	OHRank	4.43	No se especifica
	CS-LBMFL	OHRank	4.36	
[76]	Active Shape Model (ASM), Radon projection + 2D-DCT transform	SVM, SVR	4.18	EE
[77]	BIF	SVM	3.8	EE
[78]	ASM (extracción de puntos faciales), BIF	SVM, SVR	3.17	EE
			3.71	
			3.69	
[79]	ASM (extracción de puntos faciales), Filtros de Gabor, EBIF	SVM, SVR	3.17	EE
[66]	GOP, ANOVA (selección de GOP), BIF, HOG y LBP	SVM, SVR	2.81	EE
[80]	LBP	GWAgeER (GroupWise age-ranking)	2.34	EE

En la Tabla 3, se presentan los trabajos que utilizaron la base de datos MORPH para entrenamiento y evaluación. La base de datos MORPH es propuesta en [81] para entrenamiento de imágenes en ambientes controlados, puesto que las imágenes que forman ese conjunto se tomaron bajo condiciones de luz controladas, con pose de la cara al frente de la cámara y una buena resolución.

Tabla 3. Resumen y comparación de técnicas de estimación de edad sobre la base de datos MORPH.

Id	Extracción y Selección	Clasificación	MAE	Grupos de Clasificación
[68]	AAM, LBP, GABOR WAVELETS, LPQ	SVM, SVR	5.86	EE
[48]	AAM	Clustered Multi-task Support Vector Regression Machine CMTSVR	5.62	EE
[74]	Lie Algebrized Gaussians (LAG)	SVR, WCCN basado en SVM	5.21	EE
[75]	CS-LBFL	OHRank	4.52	No se especifica
	CS-LBMFL	OHRank	4.37	
[82]	BIF, Marginal Fisher Analysis MFA (reducción de dimensionalidad)	BFGS-ALDL and IIS-ALDL	4.34 ± 0.10 4.43 ± 0.05	EE
[83]	HOG, LBP and SURF	Canonical Correlation Analysis (CCA)	4.25	EE
[79]	ASM (extracción de puntos faciales), Filtros de Gabor, EBIF	SVM, SVR	4.11	EE
[77]	BIF	SVM	3.6	EE

[55]	HADF (Hierarchical Aggregation based Deep Aging Feature), Convolutional Neural Network (CNN)	SVR	3.41	EE
[66]	GOP, ANOVA (selección de GOP), BIF, HOG y LBP	SVM, SVR	2.97	EE

En la Tabla 4, se listan los trabajos encontrados en diferentes bases de datos de imágenes. El resultado reportado en la base de datos FAGE, corresponde al mejor encontrado en toda la lista de métodos, sin embargo, se debe resaltar que la base de datos se compone de pocos datos y con poca variedad de razas según reporta el autor del estimador de edad.

Tabla 4. Resumen y comparación de técnicas de estimación de edad en otras bases de datos de imágenes faciales.

Id	Extracción y Selección	Clasificación	Base de Datos	MAE	Grupos de Clasificación
[72]	Affine-invariant shape representation, AAM	SVM	Passport Dataset	9.78	EE
		Ridge Regression		15.72	
		RVM		8.84	
[51]	AAM + Local Gabor Features + selección secuencial	SVR	PAL	5.69	18-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80-89, 90-93
	AAM + LBP + selección secuencial	SVR	PAL	5.65	
[75]	CS-LBFL	OHRank	LifeSpan	5.79 5.84	No se especifica
			FACES	5.06*	
		OHRank	LifeSpan	5.26 5.62	
			FACES	4.84*	
[46]	Gabor filters, LBP, AAM	SVM, AdaBoost and LDA	BERC database,	4.6821	EE
			PAL	4.3271	
[77]	BIF	SVM	PCSO	4.1	EE
			LFW	7.8	
[83]	HOG, LBP and SURF	Canonical Correlation Analysis (CCA)	FGRC	4.17	EE
[80]	LBP	GWAgeER (GroupWise age-ranking)	FAGE (**)	1.0	EE

(*) se considera el resultado de la expresión facial neutra. (**) Base de datos local, de uso privado.

Los resultados presentados por Liu et al. en [42] muestran un porcentaje de precisión de 95.1% para una clasificación en tres grupos de edad evaluada sobre la base de datos FG-NET, y un 93,7% para una configuración en la base de datos MORPH. Sin embargo, estas medidas no son comparables con el MAE, puesto que corresponden a una estimación de edad por grupos. Dicha estimación esta sesgada no solo por las bases de datos sino también por la cantidad de grupos en los que se decide clasificar las edades, sin embargo, en búsqueda de la estandarización de la evaluación para este contexto varios autores se han regido por un protocolo de evaluación propuesto para estimación en grupos de edad [84] basado en la base de datos *Image of Groups* (IoG) [85], cuyos trabajos más destacados son presentados en la Tabla 5.

Tabla 5. Resumen y comparación de técnicas de estimación de edad en la base de datos IOG.

Id	Enfoque	AEM(%)*	AEO(%)*
[86]	Context	32.9	64.4
[87]	LaplacianScore	35.5	74.5
[86]	Appearance	38.4	71.3
[87]	FisherScore	42.8	83.7
[86]	Appearance+Context	42.9	78.1
[84]	Gabor + Adaboost	43.7	80.7
[87]	LAR	44.8	84.9
[87]	RankBoost	44.8	84.5
[84]	LBP + Adaboost	44.9	83.0
[84]	Boosted Gabor + SVM	48.4	84.4
[87]	PLO	48.5	88.0
[84]	Boosted LBP + SVM	50.3	81.1
[88]	CLBP_M + SVM	51.7	88.7
[89]	ML LBP + SVM	55.05	88.19
[90]	ML LPQ + SVM	56.0	88.8

(*) ver capítulo 4 para la definición de las métricas AEM, AEO

En las tablas 2, 3, 4 y 5, podemos apreciar evolución que ha tenido este campo de investigación en cuanto a precisión se refiere y los enfoques más relevantes encontrados en la literatura. Se destaca el uso del método AAM, dado que es utilizado frecuentemente de dos maneras: como extractor de características, y principalmente como detector facial mediante el AAM fitting. También, los filtros de Gabor demuestran ser un buen complemento para la mejora de resultados junto a otras características como la textura, apariencia y figura. De igual forma son destacados algoritmos como LBP y HOG para extraer información de textura pues obtienen buenos resultados y tienen una alta presencia en los antecedentes.

Las tablas están ordenadas de forma descendente respecto al MAE sin embargo sería incorrecto afirmar que los enfoques que tienen un MAE más bajo son los mejores porque para dar una definición de este tipo es necesario que los enfoques propuestos se evaluaran bajo un protocolo estándar, es decir sobre la misma base de datos, el mismo conjunto de imágenes para entrenamiento y test y, los mismos grupos de edad en el caso de que no se tratara sobre la estimación exacta de la edad.

2.2.2.Sistemas de Estimación Automática de Edad

Con respecto a los sistemas con estimación automática de la edad, estos emplean diferentes estrategias para tratar la información obtenida a partir del análisis facial. Una de las principales estrategias es la identificación facial que consiste en identificar clientes y registrar la información en bases de datos. Por otro lado, el análisis de video anónimo -

AVA¹⁷ implementa algoritmos de reconocimiento de patrones sin almacenar información que invada la privacidad de los clientes. A continuación, se describen algunos de estos sistemas encontrados en la literatura.

Microsoft Computer Vision API [91] permite a los desarrolladores acceder a servicios en la nube de reconocimiento facial para obtener información asociada a una imagen, por ejemplo datos demográficos. El API es accedida por un número limitado de transacciones al mes y tiene implementaciones cliente en diferentes lenguajes de programación. Una de éstas emplea REST¹⁸ para realizar llamados a través de funciones que envían un archivo de imagen o su dirección URL¹⁹ y reciben un resultado de acuerdo a la configuración especificada.

Releyeble [92] es una empresa española que implementa un sistema que calcula automáticamente la edad, sexo, emociones, detección y rastreo de miradas, entre otros. El sistema captura información mediante sensores puestos en forma estratégica sobre los establecimientos comerciales, extrae información de los clientes y la almacena en tiempo real, para ser utilizada posteriormente en procesos de inteligencia de negocios. Además, el sistema permite ofrecer campañas publicitarias a sus clientes, con base en sus características demográficas (sexo, rango de edad, entre otras).

FaceVACS-VideoScan [15] es una aplicación comercial con reconocimiento facial que realiza análisis demográfico, identificación facial, generación de estadísticas sobre métricas de audiencias²⁰, entre otras características. Este software permite en tiempo real generar información relevante sobre el comportamiento de los clientes, así como también programar eventos asociados a los clientes (por ejemplo, cuando un cliente registrado llega al establecimiento se debe informar a un asesor para darle una atención especial).

La empresa AMSCREEN [93] ofrece publicidad segmentada de productos y servicios a partir del reconocimiento automático de información demográfica capturada de sensores instalados en las pantallas que los clientes observan. Las pantallas pueden ser ubicadas dentro de los establecimientos o al aire libre, de tal manera que al acercarse los clientes, los sensores detectan sexo y edad, y el sistema muestra publicidad de acuerdo al público que observa las pantallas.

NEC FieldAnalyst [94] es un sistema que captura imágenes desde videocámaras para detectar rostros humanos y estimar automáticamente género y grupo de edad de un

¹⁷ *Anonymous video analytics* [25].

¹⁸ *Representational state transfer* [139].

¹⁹ *Uniform Resource Locator* [140].

²⁰ Métricas de audiencia: información obtenida de la observación del comportamiento de los clientes [141].

individuo para publicidad dirigida. El sistema soporta cámaras de baja resolución, como las de los circuitos cerrados de televisión, y capturan las imágenes de los clientes para su procesamiento en tiempo real.

2.2.3. Sistemas de Control Parental

Los sistemas de control parental pueden ser entendidos, como aplicaciones software que cuentan con un conjunto de funcionalidades para facilitar el monitoreo y control de las actividades realizadas, especialmente de niños, en teléfonos móviles o computadores. A continuación, serán descritas con mayor detalle este tipo de sistemas.

2.2.3.1. Sistemas de Control Parental Para Teléfonos Móviles

Estos sistemas poseen funcionalidades que permiten el monitoreo de llamadas y mensajes, conversaciones mediante aplicaciones de mensajería instantánea (chats), correo electrónico, historial de navegación, recursos multimedia, contactos, aplicaciones, calendario, contenido eliminado; rastreo mediante GPS (Global Position System), bloqueo del dispositivo de forma remota, entre otras. En la Tabla 6, son listadas las principales funcionalidades de los sistemas de control parental para dispositivos móviles disponibles en el mercado. Donde, cada columna indica la fuente de referencia que describe cada producto y en el cruce de una fila (funcionalidad) con una columna (fuente de referencia) fueron marcadas con “x” la existencia de cada una de las funcionalidades por cada producto, estas características fueron clasificadas en monitoreo y funciones.

Tabla 6. Funcionalidades de los sistemas de control parental para dispositivos móviles disponibles en el mercado.

Producto	[95]	[96]	[97]	[98]	[19]	[99]	[100]	[101]	[102]	[16]
Funcionalidad										
Monitoreo										
Llamadas y Mensajes	x	x	x	x	x	x	x	x	x	
Chats	x	x	x	x	x	x	x	x	x	
Email	x	x	x	x	x	x	x	x	x	
Historial de Navegación	x	x	x	x	x	x	x	x	x	x
Rastreo GPS	x	x	x	x	x	x	x	x	x	
Multimedia	x	x	x	x	x	x	x	x	x	x
Contactos	x	x	x	x	x	x	x	x		x
Aplicaciones	x	x	x	x	x	x	x	x	x	

Calendario	X	X	X	X	X	X	X			
Contenido Eliminado	X	X	X	X		X	X	X	X	
Keylogging	X		X	X		X	X		X	
Funciones										
Bloqueo Remoto	X	X		X	X	X	X			X
Se ejecuta de forma invisible	X		X	X	X		X		X	
Monitoreo en Tiempo Real	X	X	X				X			
	[95]	[96]	[97]	[98]	[19]	[99]	[100]	[101]	[102]	[16]

En la Tabla 6 podemos apreciar que entre las características más presentes de los sistemas de control parental para dispositivos móviles se encuentran: el control de llamadas y mensajes, chats, email, historial de navegación, monitoreo GPS, control multimedia, contactos y aplicaciones y, bloqueo remoto.

2.2.3.2. Sistemas de Control Parental Para Ordenadores

Estos sistemas se caracterizan por ofrecer funcionalidades como: bloqueo y filtro de sitios web considerados inadecuados para el usuario, bloqueo de redes sociales, bloqueo de chat, filtro de búsquedas en internet, bloqueo de transferencia de archivos, bloqueo de aplicaciones, monitoreo del historial de navegación, monitoreo de las búsquedas en internet, reportes de forma remota, monitoreo de sistemas de mensajería instantánea (chats), monitoreo del correo electrónico, entre otras. Por citar algunos ejemplos de este tipo de sistemas de control parental tenemos: Net Nanny [16], Web Watcher [102], Internet Safety [103], Witigo [101], Intego [104], Spy Agent [105], Cyber Sitter [106], Nch Software [107], Elite Keyloggers [108], cuyas funcionalidades son expuestas a continuación en la Tabla 7. Al igual que en la Tabla 6, cada columna representa las referencias a cada uno de los productos de control parental consultados y la intersección se marca con una "x" si cada funcionalidad (filas) se encuentra presente, las funcionalidades se han dividido en funciones y monitoreo.

Tabla 7. Funcionalidades de los sistemas de control parental para computadores personales disponibles en el mercado.

Producto	[16]	[102]	[103]	[101]	[104]	[105]	[106]	[107]	[108]
Funcionalidad									
Funciones									
Bloqueo y Filtro de Sitios Web	X	X	X	X	X	X	X	X	

Bloqueo de Redes Sociales	X	x	x	x	x	x	x	x	
Bloqueo de Chat	X	x	x	x	x	x	x	x	
Filtro de Búsquedas en Internet	X	x	x	x	x	x	x	x	
Bloqueo de Transferencia de Archivos	X	x	x	x	x	x	x	x	x
Bloqueo de Aplicaciones	X	x	x	x	x	x	x	x	x
Filtro de Contenido para Juegos	X	x	x	x	x				
Enmascaramiento de Profanidad	X								
Monitoreo									
Sitios Web Visitados	X	x	x	x	x	x	x	x	x
Búsquedas en Internet	X	x	x	x	x	x	x	x	x
Usuarios y Contraseñas	X	x	x	x	x	x	x		x
Redes Sociales	X	x	x	x	x	x	x		x
Reportes Remotos	X	x	x	x	x	x	x	x	x
Email	X	x	x	x		x	x		x
Chat	X	x	x		x				x
Capturas de Pantalla		x		x	x	x		x	
	[16]	[102]	[103]	[101]	[104]	[105]	[106]	[107]	[108]

Como se puede apreciar en la Tabla 7, se reconocen como las características principales en estos sistemas: el bloqueo y filtro de sitios web, redes sociales, chat, transferencia de archivos, aplicaciones, filtro de búsquedas, reporte del historial de los sitios web visitados y búsquedas en internet.

2.2.4. Sistemas de Control de Acceso y Seguridad

Los sistemas de control de acceso y de seguridad permiten el acceso a sitios o la autenticación de usuarios por medio de sistemas biométricos, a personas autorizadas y restringir (en algunos casos alertar) a personas no autorizadas. A continuación, son descritos algunos sistemas de control de acceso y seguridad que utilizan dispositivos biométricos y algoritmos de reconocimiento facial para restringir el acceso.

Los computadores portátiles Thinkpad IBM series A, T y X [109], desde el año 2000 cuentan con un sistema que monitorea continuamente quién está en frente de la terminal. Si el usuario se retira del lugar, el sistema activa el modo salva pantallas. Si un usuario que no está autorizado intenta autenticarse, el sistema niega el acceso.

El campus de La Universidad de Missouri-Rolla ha seleccionado un sistema de reconocimiento facial para asegurar un reactor nuclear [110]. El mecanismo cuenta con algunas secciones antes de llegar a un escáner facial, el cual permite regular el acceso al núcleo del reactor.

FaceGate [23] es un sistema de control de acceso que utiliza reconocimiento facial para otorgar acceso a un edificio. También se utilizan códigos asignados y tarjetas electrónicas. El sistema asigna una “clave” biométrica a la imagen facial aplicando un modelo matemático, y cada vez que un usuario desea acceder al edificio debe autenticarse con su código o tarjeta y su rostro debe coincidir con la clave almacenada.

El sistema de control de acceso biométrico estándar FaceKey [111], combina reconocimiento facial con reconocimiento dactilar para proporcionar un alto nivel de seguridad. No hay necesidad de tarjetas, claves, contraseñas o teclados. La combinación de las dos medidas biométricas hace posible tener seguridad con baja tasa de error. El sistema puede operar autónomamente, como un sistema de una sola puerta, o puede estar interconectada a múltiples puertas en múltiples sitios, nacional o internacionalmente.

FACEVACS-Entry [112] concede el acceso a un usuario si la imagen capturada por un sistema de video coincide con las características faciales almacenadas por un sistema control de acceso convencional. Para áreas de alta seguridad, es posible la combinación con pasaportes, la integración con sistemas de control existente, entre otros. Los terminales con FACEVACS-Entry pueden estar interconectados entre sí, por lo que después de la inscripción central de los datos faciales se distribuyen automáticamente a todos los terminales. El sistema puede soportar control visual para personal de seguridad.

2.3. Conclusiones Parciales

A partir del análisis y revisión de los trabajos previamente presentados, puede concluirse parcialmente que se encuentran soluciones software generalmente catalogadas como software de control parental, que mitigan el impacto de los contenidos inapropiados principalmente para los usuarios (menores de edad) de dispositivos móviles y ordenadores. Estas soluciones disponibles en el mercado se enfocan en el monitoreo y control del usuario, sin embargo, además de ser en su mayoría software propietario, ninguno manifiesta hacer uso de la estimación automática de la edad a partir de la cara para determinar de forma automática la configuración apropiada del sistema

Estas características se activan o se habilitan con un modo manual, dependiendo del administrador del sistema, generalmente un padre de familia, quien decide a cuáles usuarios o en qué momento hacer efectivas el monitoreo y control del contenido. Para automatizar dicha decisión sería idóneo integrar a esos productos, mecanismos biométricos

que hacen uso de las características faciales para el reconocimiento de la identidad y de esta forma restringir automáticamente el acceso a lugares o dispositivos.

Estos sistemas biométricos han sido acogidos en el mercado por su efectividad para soportar procesos de control de acceso a partir del reconocimiento de identidad y el cálculo automático de atributos faciales, sin embargo, a pesar de los buenos resultados obtenidos con el uso de características faciales y/o biométricas en los dominios en los que se ha probado, éstas aún no son empleadas para soportar tareas de control parental para el acceso a contenidos web desde un computador personal. Es de destacar la acogida que ha tenido este tipo de sistemas en el mercado, sin embargo, se quedaría corto en escenarios donde los dispositivos a controlar son de uso público como en escuelas, colegios, salas de internet, máquinas expendedoras, etc. debido a que se requeriría el registro de identidad de todos los usuarios en cada dispositivo.

Por otro lado, un enfoque por sistemas con capacidad de extraer información demográfica como sexo, raza y edad, mediante características faciales, es más adecuado en escenarios donde los dispositivos a controlar son de uso público y también de uso privado, entre otros motivos porque estos sistemas no persisten información sobre la identidad del usuario. En la mayoría de casos la información es usada para la generación estadísticas y reportes que soportan la toma de decisiones en establecimientos comerciales interesados en mejorar en función de las preferencias de sus clientes o para la presentación de publicidad a la medida. Estas soluciones han sido bien acogidas por el mercado y los clientes, sin embargo, al igual que los sistemas biométricos aún no son probados para soportar tareas de control parental. Cabe resaltar que no se ha encontrado evidencia relevante sobre la integración de estimadores de edad en los sistemas de control parental de contenidos web.

3. MECANISMO PROPUESTO DE ESTIMACIÓN DE LA EDAD

La estimación de la edad es considerada como un problema de reconocimiento de patrones y se ha identificado que típicamente para abordar este problema se siguen tres etapas. La primera, corresponde a la detección y normalización, la cual tiene como objetivo ubicar espacialmente la cara sobre una imagen de entrada y buscar mediante transformaciones geométricas y matemáticas una similitud en escala, rotación y tamaño en la imagen facial detectada. A partir de la imagen normalizada, en la siguiente etapa se extraen y seleccionan las características que finalmente serán usadas como entrada para así desarrollar la tercera etapa que corresponde al proceso de clasificación en el que se genera un modelo entrenado capaz de estimar la edad a partir de dichas características mediante aprendizaje supervisado. Este modelo será usado para realizar la estimación posterior al entrenamiento.

A continuación, en la Figura 7, es presentada una visión general que describe el mecanismo de estimación de la edad implementado en este trabajo.

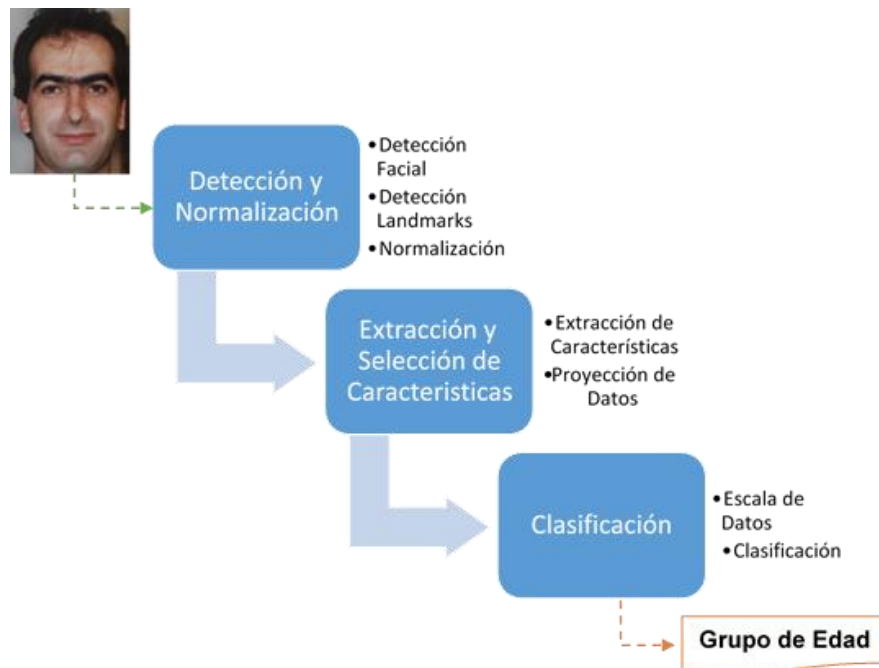


Figura 7. Vista general del mecanismo de estimación de edad.

El mecanismo propuesto consta de dos fases, una fase de entrenamiento y una de pruebas, la diferencia radica en que para el entrenamiento algunos de los procesos tienen como salida un modelo entrenado el cual es usado en la fase de prueba para realizar el proceso

de estimación. A seguir, en las siguientes subsecciones serán descritas detalladamente cada una de las etapas presentadas en la Figura 7, las cuales permitirán estimar la edad a partir una imagen facial partiendo del preprocesamiento de la misma hasta llegar a la clasificación.

3.1. Detección y Normalización

Los pasos contenidos en esta etapa de detección y normalización son presentados en la Figura 8, donde se pueden apreciar los procesos de la etapa con sus respectivas entradas, salidas y métodos usados.

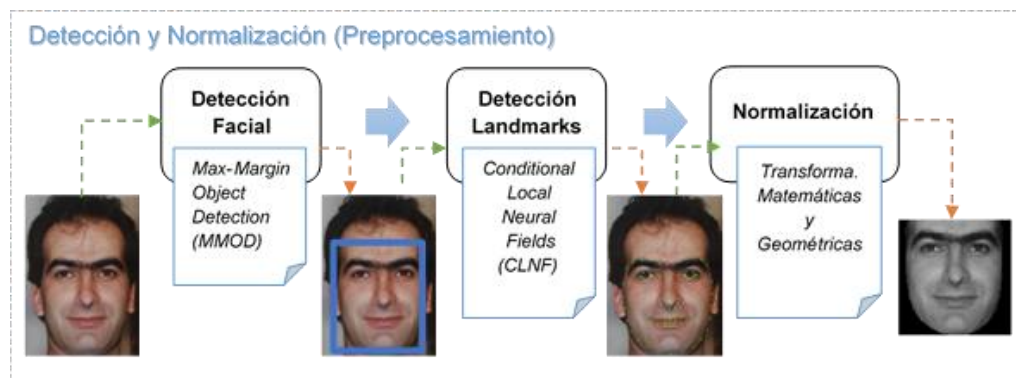


Figura 8. Vista general de la etapa de detección y normalización

En esta etapa es realizado el preprocesamiento de la imagen de entrada del mecanismo de estimación de edad y consta de tres sub-etapas: la detección facial, la detección de los landmarks y la normalización. Estas sub-etapas serán descritas a continuación en esta sección.

3.1.1. Detección facial

La detección facial, es un paso relevante durante el proceso de estimación dado que es el punto de partida para el mismo. En este trabajo fue utilizado *Max-Margin Object Detection (MMOD)*[113] para la tarea de detección facial debido a que, junto con el método de Viola Jones [114], es el de mayor presencia en los antecedentes, sin embargo, el elegido presenta una mayor precisión [113]. El algoritmo se describe así: Sea r la notación para un área rectangular de la imagen y R , el conjunto de todas las áreas rectangulares analizadas por el método de detección. Se define una etiqueta válida en una imagen como un subconjunto de R tal que cada elemento $r_i \in R$ no se solape uno con el otro. Se dice que dos rectángulos r_1 y r_2 no se solapan si la proporción de su área de intersección al área total cubierta es inferior a 0,5 como lo describe la ecuación (10).

$$\frac{Area(r_1 \cap r_2)}{Area(r_1 \cup r_2)} < 0.5 \quad (10)$$

En la ecuación (11), a partir de una imagen x y una función de puntuación de ventana (*window scoring function*) f , se puede definir el proceso de detección.

$$y^* = \underset{y \in Y}{\operatorname{argmax}} \sum_{r \in y} f(x, r) \quad (11)$$

Donde Y representa el conjunto de todas las etiquetas válidas. Se pretende encontrar el conjunto de las posiciones de *window sliding* que tienen las puntuaciones más grandes, pero al mismo tiempo que no se superponen. Esto se logra típicamente con el método de clasificación de “pico codicioso” (*greedy peak sorting*), el cual a partir de un conjunto de rectángulos r tales que $f(x, r) > 0$ para todo $r \in R$, se agregan a un conjunto resultado y^* uno a uno de mayor a menor según su puntuación de ventana, siempre y cuando el rectángulo a agregar no se sobreponga con los que ya están en el conjunto y^* , en consecuencia el rectángulo con mayor puntaje de ventana tendrá prioridad sobre los demás, lo que puede resultar en una solución no óptima (Figura 9), por ende la ecuación (11) no es precisamente factible desde el punto de vista computacional. En la Figura 9, se ejemplifica, asumiendo solapamiento, es decir que no cumplen con la ecuación (11). Una solución óptima debería seleccionar r_1 y r_3 con una puntuación total de 10, sin embargo, mediante el algoritmo de clasificación de pico codicioso selecciona r_2 con una puntuación total de 6.

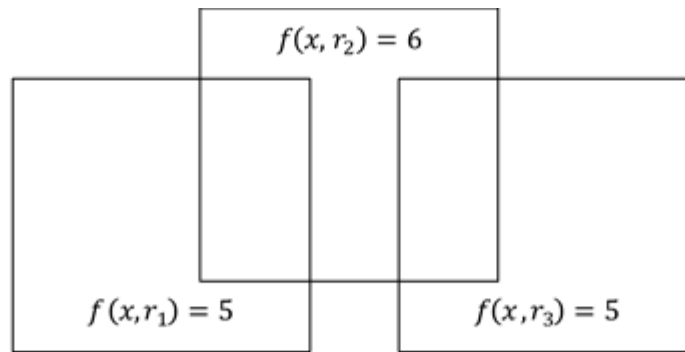


Figura 9. Tres *window sliding* y sus respectivos puntajes

Idealmente un algoritmo de aprendizaje se encargaría de encontrar la función de puntuación de ventana que minimiza conjuntamente el número de falsas alarmas y detecciones omitidas por el algoritmo de clasificación de pico codicioso. MMOD considera solamente funciones de puntuación de ventana que tienen la forma:

$$f(x, r) = \langle \omega, \phi(x, r) \rangle \quad (12)$$

Donde ϕ refiere a la función de extracción de características de la *sliding window* en la ubicación r sobre la imagen x , y ω es un vector de parámetros. Así reemplazando (12) en (11), tenemos:

$$y^* = \underset{y \in Y}{\operatorname{argmax}} \sum_{r \in y} \langle \omega, \phi(x, r) \rangle \quad (13)$$

Si hacemos:

$$F(x, y) = \sum_{r \in y} \langle \omega, \phi(x, r) \rangle \quad (14)$$

Entonces a partir de (13) tenemos:

$$y^* = \underset{y \in Y}{\operatorname{argmax}} F(x, y) \quad (15)$$

El objetivo es buscar un vector ω de parámetros que conduzca a minimizar el error en las detecciones, búsqueda que se lleva a cabo mediante una selección aleatoria de pares imagen, etiqueta $(x_i, y_i) \in X \times Y$, donde la premisa es que la puntuación para el etiquetado correcto de x_i sea mayor que la puntuación de las etiquetadas erróneamente. Así la ecuación (16) debe satisfacerse tan a menudo como sea posible.

$$F(x_i, y_i) > \max_{y \neq y_i} F(x_i, y) \quad (16)$$

Como resultado de este proceso tenemos la ubicación de cero, uno o varios rectángulos que corresponden cada uno a la ubicación de una cara dentro de la imagen de entrada del mecanismo. Esta información es usada por el siguiente proceso, detección de landmarks, como entrada para cumplir con su objetivo. Se define esta etapa como el punto de partida pues no tiene sentido continuar el proceso de estimación de la edad si el proceso de detección facial arroja un resultado de cero (0) caras en la imagen.

3.1.2. Detección de puntos de referencia (*landmarks*)

Este proceso tiene como objetivo ubicar dentro del rectángulo de la cara detectada, el contorno y la ubicación de ojos nariz y boca. El mecanismo de estimación será usado en un ambiente real, por ende, debe ser capaz de detectar los landmarks y realizar seguimiento de las caras a partir de las imágenes de entrada en condiciones de ambiente no controladas, es decir, con condiciones bajas de luz y presencia de oclusión. *Conditional Local Neural Fields* (CLNF) [115] es una extensión del método *Constrained Local Model* (CLM) [116] que se enfoca en la detección de características en ambientes complejos superando a sus predecesores en cuanto a la precisión de la detección de *landmarks* en dichos ambientes.

CLM usa como parámetro de entrada la ubicación de la cara obtenida previamente y consta principalmente de tres partes: *point distribution model* (PDM), detectores locales y el método de ajuste (*fitting*). PMD es usado para la ubicación de los puntos de las características faciales en las imágenes. Cada ubicación es calculada mediante la ecuación (17)

$$x_i = s * R * (\bar{x}_i + \Phi_i q) + t \quad (17)$$

Donde x_i representa la ubicación de la i -ésima característica, y $p = \{s, R, t, q\}$ representa los parámetros de PDM: s es la escala global, R la rotación, t la traslación y q el conjunto de parámetros no rígidos. \bar{x}_i representa el valor medio de la i -ésima característica (por ejemplo: $\bar{x}_i = [\bar{x}_i, \bar{y}_i]$ para un modelo 2D). Φ_i representa una matriz de componentes principales.

La búsqueda de los parámetros p se realiza mediante la ecuación:

$$p^* = \underset{p}{\operatorname{argmin}} [R(p) + \sum_{i=1}^n D_i(x_i; I)] \quad (18)$$

Donde R representa el termino de regularización que penaliza las formas demasiado complejas o poco probables y D la medida de desalineación que el i -ésimo *landmark* experimenta en la posición x_i sobre la imagen I .

Otra de las partes del CLM corresponde a los detectores locales que se encargan de evaluar la probabilidad de que un *landmark* se encuentre alineado con la ubicación de determinados pixeles con el fin de capturar las variaciones de apariencia local en cada landmark. Cada detector local π_{x_i} en la ubicación x_i de la imagen se define como

$$\pi_{x_i} = C_i(x_i; I) \quad (19)$$

La desalineación puede modelarse usando un regresor cuya salida indique un valor entre 0 y 1, donde 0 indica la no alineación y 1, una alineación perfecta. Así C_i corresponde a la salida del regresor para la i -ésima característica. En esta implementación para los detectores locales se utiliza el *método Local Neural Field* (LNF). Este método puede capturar las relaciones entre pixeles tanto vecinos como distantes, aprendiendo tanto de la similitud como de las restricciones de dispersión de larga distancia. Además, puede capturar relaciones complejas no lineales entre valores de pixeles y las respuestas de salida mediante la inclusión de una capa de redes neuronales (*neural network layer*) [115]. En la Figura 10 se muestra como el resultado obtenido (Figura 10 c) es una buena aproximación al resultado ideal (Figura 10 b)

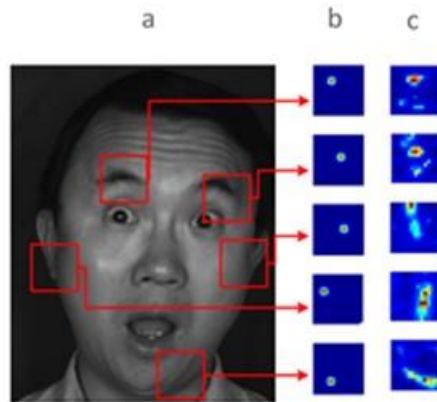


Figura 10. Ejemplo de aplicación de método de detección de puntos de referencia
(a) Características ejemplo. (b) Resultado ideal. (c) Resultado al aplicar LNF
(Adaptado de CLNF [115])

La tercera parte del CLM corresponde al método de ajuste (*fitting*). CLM es un enfoque local y se basa en una estimación inicial de parámetros proporcionada por el detector facial. Teniendo un \mathbf{p}_0 el objetivo es encontrar un $\Delta\mathbf{p}$ que aproxime a una solución $\mathbf{p}^* = \mathbf{p}_0 + \Delta\mathbf{p}$ donde \mathbf{p}^* es una solución óptima. Así el objetivo de ajuste (*fitting*) es presentado en la ecuación (20)

$$\underset{\Delta\mathbf{p}}{\operatorname{argmin}}[R(\mathbf{p}_0 + \Delta\mathbf{p}) + \sum_{i=1}^n D_i(x_i; \mathbf{I})] \quad (20)$$

Este puede ser resuelto de múltiples formas. El más común refiere al método *Regularised Landmark Mean Shift (RLMS)*, sin embargo en esta implementación se utiliza la mejora que presenta Baltrusaitis y Robinson [115] denominada *Non-Uniform Regularised Landmark Mean Shift (Non-Uniform RLMS)* que toma en cuenta la fiabilidad de cada descriptor local a diferencia de RLMS donde cada descriptor local es igualmente fiable.

3.1.3. Normalización

A partir de la información extraída con la detección facial y la detección de landmarks, cada imagen de entrada es normalizada con el fin de que todas las imágenes estén en las mismas condiciones de rotación de la cara, escala y espacio de color. Sin importar la rotación que tenga la cara en la imagen de entrada, mediante la información de los landmarks la cara es rotada verticalmente. En cuanto a la escala, las imágenes son ajustadas a un tamaño de 96 píxeles de alto por 96 píxeles de ancho. Se hace uso de un espacio de color, las imágenes son pasadas a escala de grises, además los valores de gris diferentes a 0 solo son asignados en la imagen normalizada si cada pixel de la imagen corresponde a un pixel dentro de la cara, lo que resulta en una imagen normalizada cuyos

valores de gris fuera de la cara son 0, de esta forma la cara es recortada y se descartan píxeles en la imagen que no son de interés para el proceso de extracción de características.

Para la detección y la normalización se ha hecho uso del framework de código abierto Open Face [117]. Este framework implementa métodos para el análisis facial, entre sus funcionalidades presenta detección facial, de landmarks y seguimiento, estimación de pose de la cabeza, reconocimiento de AU (*Action Unit*) faciales y estimación de la mirada.

Una característica importante para nuestra propuesta se refiere al funcionamiento en tiempo real, lo cual para el objetivo de este trabajo es una condición ideal, dado que el control parental incluye tareas de monitoreo y control en tiempo real. Para la implementación de lo expuesto anteriormente fue utilizada la librería OpenFace que a su vez integra la librería dlib[113] para la detección facial mediante MMOD. Además, incluye el método CLNF para la detección de los *landmarks* y proporciona un modelo ampliamente entrenado para ello. También entre sus funcionalidades ofrece la alineación facial que permite hacer el proceso de normalización. A pesar del abanico de funcionalidades que ofrece el framework para el análisis facial, en este trabajo hemos usado solo las relacionadas con la detección facial, detección de *landmarks* y alineación de la cara.

Un ejemplo del proceso de detección y normalización descrito puede verse en la Figura 11, en la cual a partir de la imagen de entrada (a) se realiza la detección facial representada en (b), la cual a su vez es usada como parámetro para la detección de los landmarks (c) que contribuye a la normalización e imagen resultado del proceso (d).

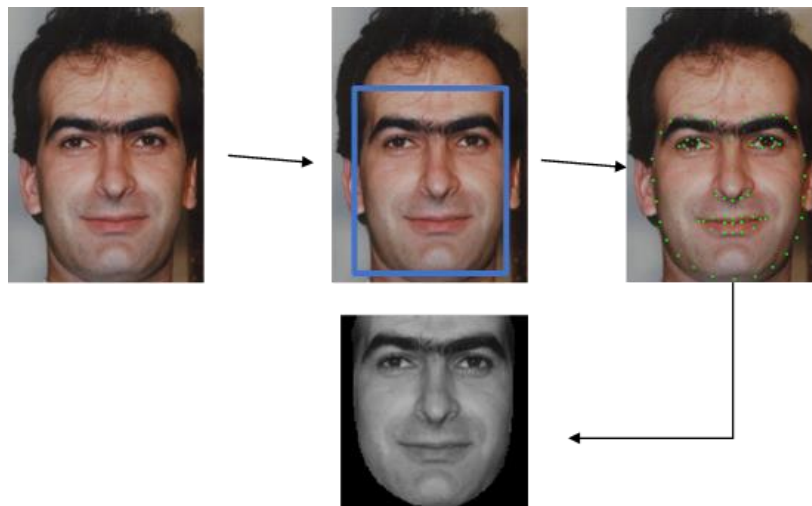


Figura 11. Ejemplo del proceso completo de detección y normalización. (a) Imagen de entrada. (b) Detección facial. (c) Detección de landmarks. (d) Imagen normalizada.

Como se describe a lo largo de esta sección, la etapa de detección y normalización proporciona como resultado una imagen facial con características similares en escala,

rotación y espacio de color para cada imagen de entrada del sistema de estimación, también reduce el ruido de la imagen de entrada, es decir, descarta píxeles de la imagen que no proporcionan información relevante para el proceso de estimación. Este resultado es la entrada de la siguiente etapa que refiere a la extracción y selección de características.

3.2. Extracción de características

En esta sección se describe la etapa de extracción de características, la Figura 12, presenta la vista general de esta etapa, donde pueden apreciarse los pasos con sus respectivas entradas, salidas y métodos usados. La entrada de esta etapa es una imagen facial normalizada, que es referencia para la generación de las salidas que corresponden al modelo de entrenamiento de proyección de datos, para la fase de entrenamiento del mecanismo y, la proyección de las características extraídas de la imagen normalizada que funciona también como entrada de la siguiente etapa tanto para la fase entrenamiento como para la de prueba.

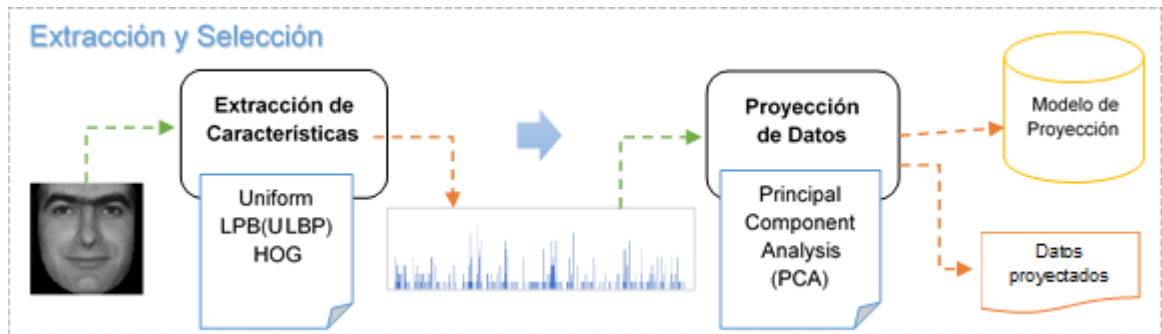


Figura 12. Vista general de la etapa de extracción y selección de características

Para la implementación del proceso de extracción de características fue identificado el algoritmo de *Local Binary Pattern* (LBP) [118] como un extractor de características locales ampliamente usado para la descripción de texturas en el campo del reconocimiento de patrones. Además, en el campo de estimación de la edad es el método de extracción de características más referenciado durante el levantamiento del estado del arte. LBP es simple, eficiente y robusto frente a variaciones en la intensidad del nivel de gris derivadas, entre otras cosas, de las diferencias en la iluminación.

En general LBP recorre cada pixel de la imagen y lo toma como pixel central de un vecindario, su valor de gris es comparado con el de sus vecinos y dicha comparación es codificada mediante un número binario el cual es asignado al pixel central. Formalmente LBP se describe según la ecuación (21) y (22)

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad (21)$$

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (22)$$

Para la ecuación (21), P es el número de vecinos que se comparan, R es el tamaño del vecindario y, g_c y g_p son los valores de gris del pixel central y cada uno de los p pixeles del vecindario respectivamente teniendo $p = 0 \dots P - 1$. En la Figura 13 podemos apreciar el cálculo del LBP sobre un pixel para los parámetros P = 8, R =1, en este ejemplo el código LBP es 10000001 (127 en decimal).

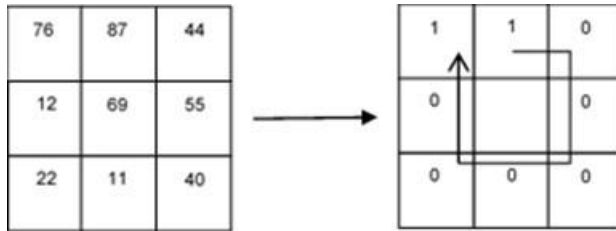


Figura 13. Ejemplo del Operador LBP básico

La técnica LBP descarta la información sobre la magnitud de la diferencia entre el valor de gris del pixel central y su vecindario, ya que solo considera el signo de dicha diferencia, pues el valor del código LBP solo depende del resultado de la comparación, lo que puede ocasionar la generación de códigos inconsistentes como lo podemos apreciar en la Figura 14. En este ejemplo mediante el proceso de codificación se genera un código LPB uniforme [1,1,1,1,1,1,1,1] que corresponde a un área plana o un punto oscuro [119] lo cual no es correcto.

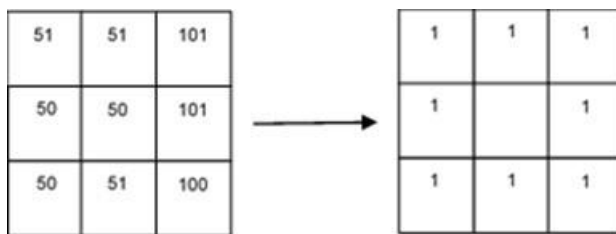


Figura 14. Generación de patrón binario inconsistente en el proceso de codificación LBP

Para mitigar este problema presentado por este mecanismo, se ha contemplado una extensión de LBP denominada *Compound Local Binary Pattern* (CLBP) usada en otros problemas de reconocimiento de patrones como el reconocimiento de expresiones faciales [120] y destacada en el campo de estimación del género, raza y edad [24]. A diferencia de LBP que emplea un bit por cada vecino para representar el signo de la diferencia entre el valor de gris del pixel central y el del vecino. CLBP emplea un código con longitud de 2P-bits, es decir 2 bits por vecino, donde el primer bit corresponde al valor de LBP original y el

segundo representa la magnitud de dicha diferencia. Formalmente CLBP se define según la ecuación (23) y (24)

$$CLBP_{P,R} = \sum_{p=0}^{P-1} u(g_p, g_c) 2^p \quad (23)$$

$$u(g_p, g_c) = \begin{cases} 00 & g_p - g_c < 0, |g_p - g_c| \leq M_{avg} \\ 01 & g_p - g_c < 0, |g_p - g_c| > M_{avg} \\ & g_p - g_c \geq 0, |g_p - g_c| \leq M_{avg} \\ 11 & \text{en otro caso} \end{cases} \quad (24)$$

Para la ecuación (23), P es el número de vecinos que se comparan, R es el tamaño del vecindario y, g_c y g_p son los valores de gris del pixel central y cada uno de los p pixeles del vecindario respectivamente teniendo $p = 0 \dots P - 1$.

En la ecuación (24), M_{avg} es el promedio de la magnitud de la diferencia entre los valores de gris de los vecinos y el pixel central. Para la implementación consideramos un número de vecinos de 8 ($P = 8$) alrededor del pixel central con un tamaño del vecindario de 1 ($R = 1$) lo que da como resultado al aplicar CLBP sobre un pixel, un código de 16 bits que provee 8 bits con la información de magnitud y 8 bits del signo. Este código es separado en dos sub-CLBP patrones obtenidos mediante la concatenación de los valores correspondiente en la secuencia de bits (1,2,5,6... 2P-3, 2P-2) y (3,4,7,8... 2P-1, 2P) del código CLBP original de tamaño 2P-bits (16 bits) como se muestra en la Figura 15.

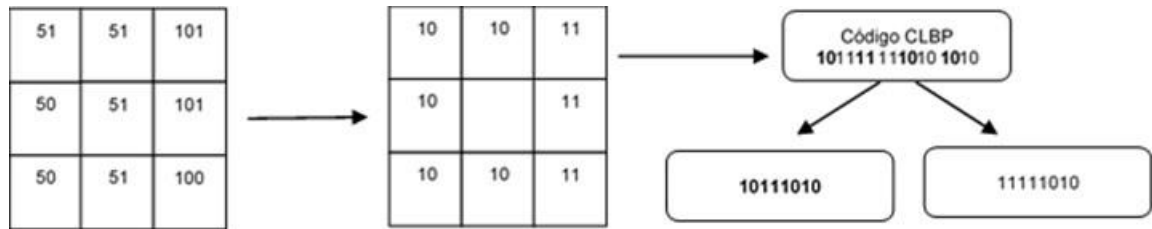


Figura 15. Operador CLBP y cálculo de los sub-CLBP patrones.

De este modo al aplicar CLBP a una imagen, se obtienen dos imágenes codificadas, una por cada sub-CLBP. En consecuencia, por cada imagen se obtienen dos histogramas, los cuales se concatenan para formar el histograma CLBP. Con el propósito de tener en cuenta la información espacial [121] se ha dividido cada imagen de la cara normalizada en sub regiones no solapadas. La imagen es dividida en 5 filas y 5 columnas formando una matriz de 25 celdas. Como se describe en el Capítulo 5, se han implementado y evaluado tanto CLBP como LBP, cada descriptor se aplica a cada celda de forma separada e independiente y los histogramas resultantes de cada celda se concatenan uno a uno para formar un histograma general que describe toda la imagen, este proceso descrito puede observarse en la Figura 16.

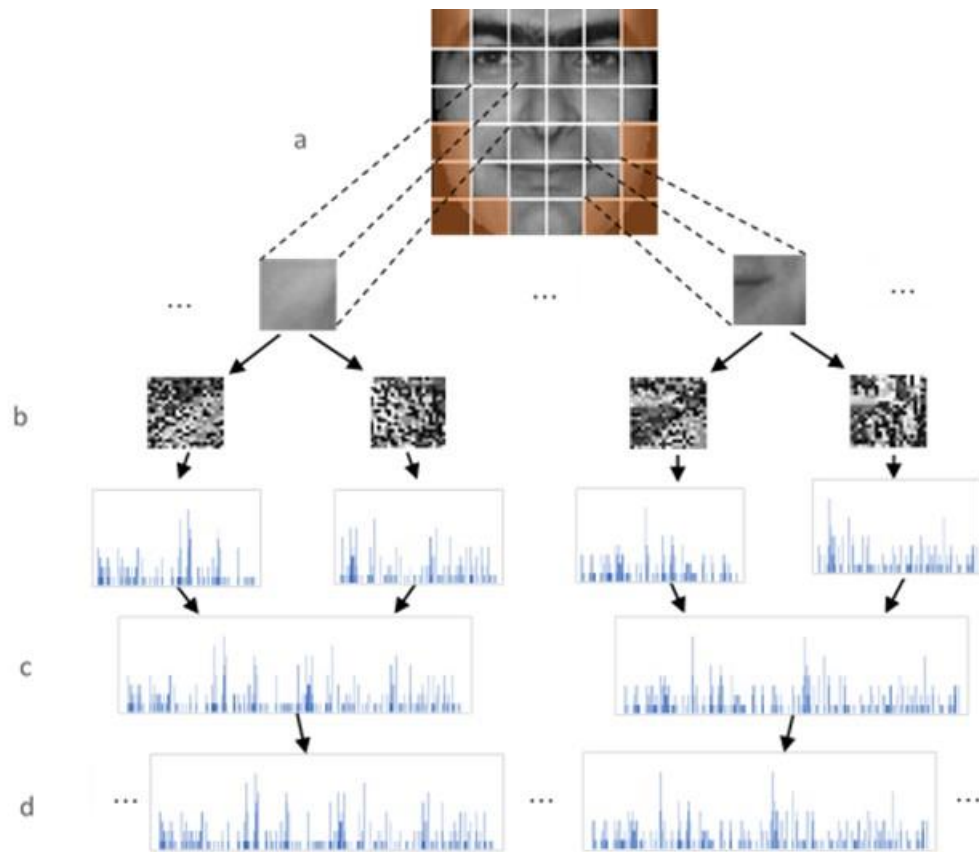


Figura 16. Proceso de extracción de características con CLBP. (a) División en celdas de la imagen. (b) Extracción de los códigos CLBP. (c) Histograma celda a celda. (d) Histograma general de la imagen

En esta implementación se han omitido las celdas marcadas sobre la matriz que divide la imagen pues no tienen información relevante en cuanto a textura de la cara, en consecuencia, son irrelevantes para la estimación de la edad. Cada celda es recortada y tratada de forma independiente, es decir, a cada celda se le ha aplicado CLBP y LBP individualmente de tal forma que se obtienen dos imágenes codificadas por cada celda en el caso de CLBP y una para LBP. Por ende, tendríamos un vector de características de 512 por celda en el primer caso y 256 en el segundo, pero con el fin de reducir el número de características se aplica *Uniform Local Binary Pattern* (ULBP)[122], método que permite la representación del histograma en 59 características usando lo que se conoce como patrones uniformes, por lo que en total tendríamos 118 características por cada celda para CLBP y 59 para LBP. Los histogramas de las imágenes codificadas son calculados y concatenados para obtener un histograma general de la celda. Se repite el proceso por cada celda (26 veces en total) y los histogramas se concatenan uno tras de otro para formar un histograma de toda la imagen.

Para complementar la información de textura otorgada por los descriptores anteriormente mencionados, se ha implementado un segundo descriptor, el *Histogram of Oriented Gradients* (HOG) que ha sido ampliamente usado en problemas de visión artificial relacionadas con la detección de objetos y, de igual forma en el contexto de la estimación de la edad.

En primera instancia existe la necesidad de definir el gradiente como un cambio de la intensidad de la imagen en la dirección donde el cambio de intensidad es máximo, se calcula para cada uno de los pixeles de la imagen y se define mediante dos valores, la dirección del cambio de intensidad máximo y la magnitud del cambio en esa dirección. El cálculo del gradiente se hace basado en la diferencia de intensidad de los pixeles adyacentes en dirección horizontal y vertical:

$$dx = I(x + 1, y) - I(x - 1, y) \quad (25)$$

$$dy = I(x, y + 1) - I(x, y - 1) \quad (26)$$

Donde I refiere al valor del pixel en dicha posición.

Como puede apreciarse en la imagen Figura 17, al aplicar la formula podríamos trasladar su resultado a un eje de coordenadas los cuales nos definen un vector gradiente para ese pixel a partir del cual podemos calcular la magnitud y la dirección del gradiente mediante las ecuaciones (28) y (27) respectivamente, en ese y cada uno de los pixeles de la imagen:

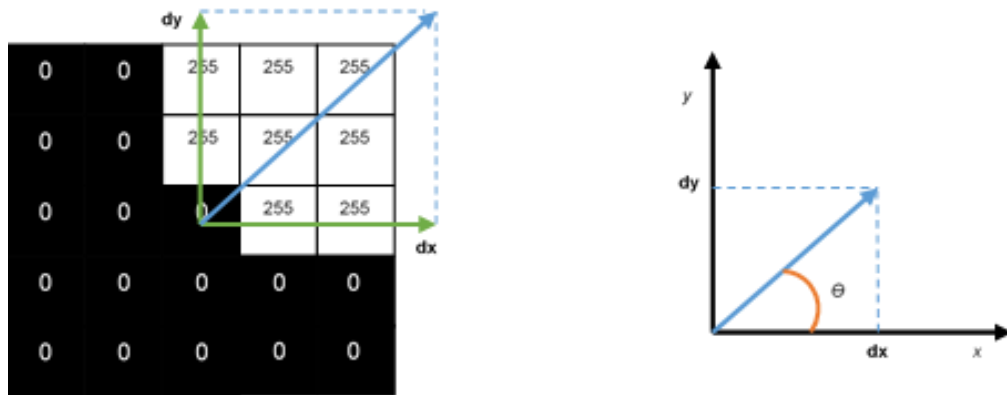


Figura 17. Calculo de gradiente

$$\theta(x, y) = \arctan\left(\frac{dy}{dx}\right) \quad (27)$$

$$g(x, y) = \sqrt{dx^2 + dy^2} \quad (28)$$

Para conseguir la descripción global de la imagen mediante el cálculo de HOG, como se presenta en la Figura 18, en primer lugar, se debe dividir la imagen a analizar en un tamaño fijo de celdas y para cada una de las celdas se obtiene un histograma de las orientaciones

de los gradientes en esa celda. Para todas y cada una de las celdas se repite el proceso y en seguida se combinan todos los histogramas para obtener una representación global de la imagen en forma de vector de características. Esto con el fin de obtener las orientaciones dominantes en la imagen, así como la información espacial sobre la distribución de los gradientes.



Figura 18. Proceso extracción HOG

Para calcular los histogramas de orientación en una celda debemos elegir una serie de parámetros: el tamaño de la celda en píxeles, el rango de las orientaciones que puede ser 180° o 360° y el número de intervalos en los que se divide dicho rango. En este trabajo se escogieron celdas de 8 píxeles de alto y ancho, un rango de orientación de 180° lo cual implica que los gradientes que tengan la misma orientación, pero sentido diferente, serán considerados equivalentes y por ende se asignarán a un mismo intervalo; por último, el número de intervalos en los que dividiremos el rango será 9, por lo que cada rango tendrá un intervalo $\delta\theta$ de 20° (Figura 19).

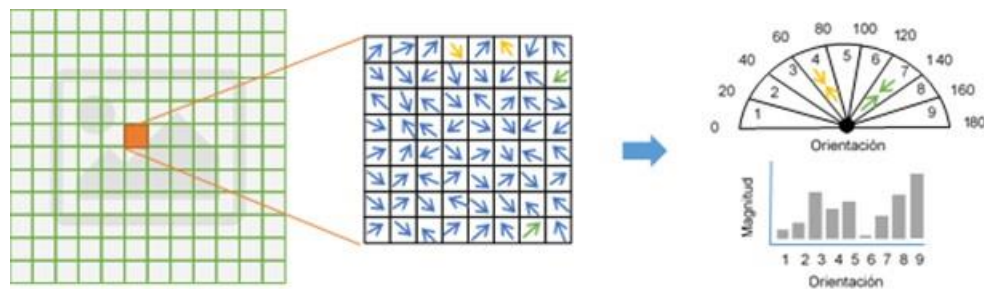


Figura 19. Cálculo del histograma de gradientes.

Una vez fijados estos parámetros, cada uno de los gradientes de la celda con su orientación y magnitud quedará asignado a uno de los intervalos en función de la orientación del gradiente. De esa forma a cada intervalo se le asignarán los gradientes cuya orientación esté dentro de los límites designados. Finalmente, el valor de cada intervalo en el histograma resultante se obtiene al sumar la magnitud de sus gradientes asignados. Formalmente el valor del histograma para una orientación k , donde $1 \leq k \leq 9$ se define como:

$$h(k) = \sum_{(x,y) \in C} \omega_k(x,y)g(x,y) \quad (29)$$

Donde $h(k)$ representa la magnitud del histograma para la orientación k , C es la celda sobre la que se está calculando el histograma, $g(x,y)$ corresponde a la magnitud del vector gradiente y $\omega_k(x,y)$ es el factor que determina la asociación del gradiente sobre el intervalo k e inicialmente es definida como sigue

$$\omega_k(x,y) = \begin{cases} 1, & \text{si } (k-1)\delta\theta \leq \theta(x,y) < k\delta\theta \\ 0, & \text{en el caso contrario} \end{cases} \quad (30)$$

Donde $\delta\theta$ es el tamaño del intervalo en los que se divide el rango de orientación, 20 para este caso. De esta forma se asigna cada gradiente a un único intervalo, sin embargo, con este modelo gradientes con orientaciones muy similares pueden asignarse a intervalos diferentes (Figura 20), por lo que se debe asignar cada gradiente a los dos intervalos más cercanos con un peso proporcional a la distancia de la orientación al centro de cada intervalo, de esta forma el factor de asociación está dado por la ecuación (31)

$$\omega_k(x,y) = \max\left(0, 1 - \frac{\theta(x,y) - \theta_k}{\delta\theta}\right) \quad (31)$$

Donde $\theta(x,y)$ es la orientación del gradiente, θ_k el valor medio del k intervalo de orientación y $\theta(x,y) - \theta_k$ la distancia entre ambos. Es de notar que para distancias mayores a las de $\delta\theta$, que corresponde al intervalo de orientación, el segundo término de la ecuación tomará valores negativos por lo que el factor de asignación será 0.

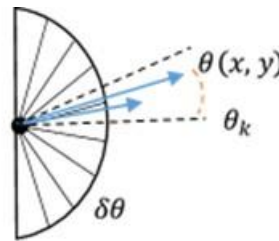


Figura 20. Vectores de Gradiente con orientaciones similares asignados a intervalos de orientación diferentes.

De igual forma bajo el esquema que cada gradiente únicamente contribuye a la celda que pertenece puede presentar problemas similares a los que se contemplaron con los histogramas de orientaciones similares. Un ejemplo lo podemos apreciar en la Figura 21. donde dos pixeles muy cercanos en la imagen pueden quedar en celdas diferentes por lo que pequeños cambios pueden resultar en variaciones significativas en la representación final. Para solucionar este problema, de forma análoga al problema de la orientación, se debe asignar cada gradiente a las cuatro celdas más cercanas con un peso proporcional a la distancia del pixel al centro de cada celda, de esta forma el factor de asociación está dado por la ecuación (32) en la dirección horizontal y (33) para la vertical

$$\omega_{ij}^x(x, y) = \max(0, 1 - \frac{d_{ij}^x}{\delta x})(32)$$

$$\omega_{ij}^y(x, y) = \max(0, 1 - \frac{d_{ij}^y}{\delta y})(33)$$

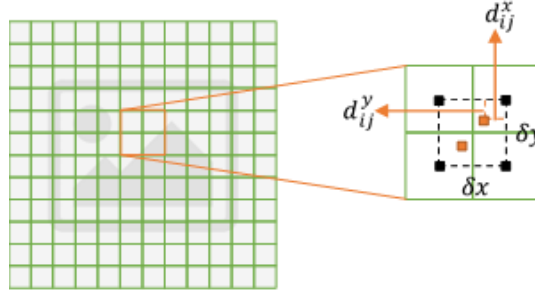


Figura 21. Pixeles cercanos asignados a celdas diferentes.

Donde ij refieren a la posición del pixel en la celda, d_{ij}^x y d_{ij}^y corresponden a la distancia desde el pixel hasta el centro de la celda y, δx y δy es la distancia entre los centros de las celdas. En este sentido el valor final del histograma esta dado por la ecuación (34)

$$h_{ij}(k) = \sum_{(x,y)} \omega_{ij}^x(x, y) \omega_{ij}^y(x, y) \omega_k(x, y) g(x, y) \quad (34)$$

En seguida de calcular el valor del histograma para cada una de las celdas en las que fue dividida la imagen, existe la necesidad de normalizar el resultado para mitigar el impacto en los valores del histograma causados por los cambios de iluminación. Las celdas con un contraste alto presentarán magnitudes de gradiente más altos que las de contraste bajo. Sin embargo, los cambios de iluminación pueden o no ser constantes a lo largo de la imagen por lo que no es suficiente aplicar una sola normalización par toda la imagen, por el contrario, es más conveniente una normalización local que para este caso será definida por bloques. Cada bloque contendrá un número $l * l$ de celdas, en este trabajo $l = 2$, por lo que cada bloque estará compuesto por 4 celdas. Para realizar la normalización, se concatenan los histogramas de las celdas del bloque para obtener un vector v con la representación del bloque el cual se normalizará mediante la ecuación (35)

$$v' = \frac{v}{\sqrt{\|v\|_2^2 + \epsilon}} \quad (35)$$

Teniendo que:

$$\|v\|_2 = \sqrt{\sum x_i^2} \quad (36)$$

Los bloques se definen de forma que tengan solapamiento entre ellos, la redundancia que otorga el solapamiento ayuda a obtener un descriptor más robusto. Cada bloque se ubica

con una separación de una celda tanto horizontal como verticalmente (Figura 22 a) y la representación final del descriptor HOG se obtienen concatenando la representación normalizada de cada uno de los bloques solapados (Figura 22 b).

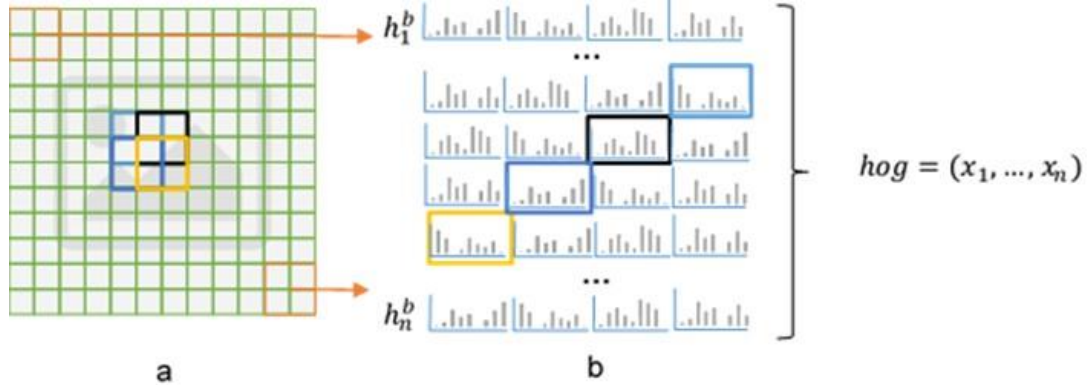


Figura 22. (a) Definición de los bloques de celdas (b) Contribución de una celda al vecindario de bloques, normalización y vector de características.

Debido a este solapamiento cada celda en la que se divide la imagen contribuye en la descripción de varios bloques, pero en cada uno de forma diferente pues la normalización de cada bloque se calcula independientemente. Así como se puede apreciar en la Figura 22 donde se puede observar la contribución de una misma celda a la descripción de cuatro bloques diferentes.

Posteriormente, el número n de dimensiones o características extraído por el descriptor HOG está dado por la ecuación (37)

$$n = n^{\circ}bloques * n^{\circ}celdas|bloque * n^{\circ}intervalos_histograma \quad (37)$$

Teniendo

$$n^{\circ}bloques = n^{\circ}celdas - n^{\circ}celdas|bloque + 1 \quad (38)$$

Después de la extracción de los vectores de características mediante los descriptores CLBP o LBP y HOG, estos se concatenan y se normalizan en el rango $\{0,1\}$.

Con el fin de reducir su dimensionalidad, el análisis de componentes principales (PCA) es utilizado para reducir el número de características de los datos extraídos por los descriptores [70]. El resultado es un número menor de variables denominadas *componentes principales*, éstas son una combinación lineal de las variables iniciales y sintetizan la mayor parte de la información contenida en los datos originales, independientes entre sí. Un aspecto importante del PCA es el análisis de la matriz de correlaciones, puesto que una correlación alta entre variables es indicativa que existe información redundante. Los componentes principales se escogen de aquellos factores cuyo porcentaje de

variabilidad se considere suficiente. De esta forma, se representa una matriz a partir de los componentes principales donde cada celda muestra las correlaciones entre las variables y los componentes principales. El número de componentes principales equivale al número de columnas de la matriz y las filas son las variables.

En este trabajo fue implementado para la reducción de características, el algoritmo PCA propuesto por Belhumeur et al. [123], usada en un amplio grupo de trabajos [24] [46] [124] en el contexto de estimación automática de edad con características extraídas con descriptores de textura. Realizar dicha reducción de dimensionalidad es un paso clave en cuanto a términos de complejidad computacional pues hace que la carga de trabajo para el SVM sea menor, además este algoritmo atiende una de las etapas contempladas por el mecanismo a proponer.

3.3. Clasificación

Después de la reducción de dimensionalidad, la proyección de datos es usada como entrada para el algoritmo de clasificación. En la Figura 23 es presentada la vista general de la etapa de clasificación donde pueden observarse los procesos de esta etapa con sus respectivas entradas, salidas y métodos usados. Como resultado de esta etapa se obtiene el modelo SVM generado para la fase de entrenamiento y como salida final de la etapa y el mecanismo, la predicción del grupo de edad al que pertenece la persona retratada en la imagen facial de entrada.

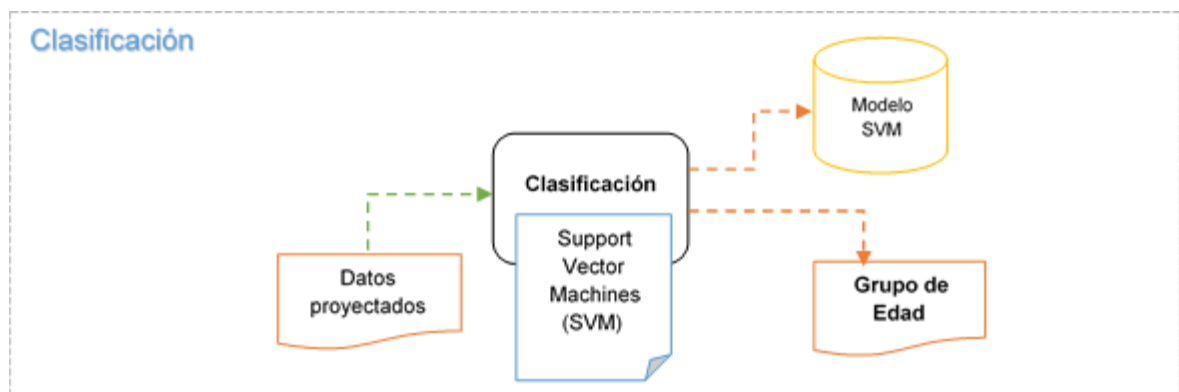


Figura 23. Vista general de la etapa de clasificación

El proceso de clasificación fue realizado mediante el uso de *Support Vector Machines* (SVM) [33], el cual es un algoritmo de clasificación supervisado y el método más usado como algoritmo de clasificación en el estado del arte para el problema de la estimación automática de la edad. Un conjunto de datos de entrenamiento es etiquetado y tomado como argumento de entrada, así el algoritmo de clasificación obtiene una frontera de decisión óptima para separar los datos basados en sus etiquetas correspondientes. El

resultado de este proceso es un modelo de predicción de las etiquetas de datos de prueba no incluidos como datos de entrenamiento.

SVM fue implementado mediante el uso de la librería de código abierto Libsvm [125]. Esta librería ofrece una serie de formulaciones de SVM para clasificación, regresión y estimación de distribución. Las formulaciones integradas son: *C-support vector classification* (C-SVC), *v-support vector classification* (v-SVC), *distribution estimation (one-class SVM)*, *ε-support vector regression* (ε-SVR), y *v-support vector regression* (v-SVR).

En este trabajo se hace uso de C-SVC para clasificación de múltiples clases. El objetivo del método es encontrar la solución al problema de optimización de la ecuación (39), teniendo los vectores de entrenamiento $x_i \in R^n, i = 1 \dots l$ y el vector indicador $y \in R^l$ tal que $y_i \in \{1, -1\}$. Es decir, las parejas ítem – valor $(x_i, y_i), i = 1 \dots l$.

$$\min_{w,b,\varepsilon} \frac{1}{2} w^T w + C \sum_{i=1}^l \varepsilon_i \quad (39)$$

$$\text{Sujeto a: } y_i(w^T \phi(x_i) + b) \geq 1 - \varepsilon_i, \varepsilon_i \geq 0, i = 1 \dots l \quad (40)$$

Los vectores de entrenamiento x_i son mapeados en un espacio altamente dimensional mediante la función ϕ . $C > 0$ es el parámetro de regularización. Para resolver problemas linealmente no separables o la clasificación en más de dos clases, como en este caso, se hace uso de la función *kernel Radial Basis Function* (RBF) descrita en la ecuación (41) como se sugiere para problemas linealmente no separables[126], y también usada en propuestas anteriores [124] [127][88]

$$K(x, y) = e^{-\gamma \|x-y\|^2} \quad (41)$$

Con los datos proyectados y la función *kernel* definida, se procede a buscar los parámetros γ y C , como serán definidos posteriormente durante el proceso de evaluación en el Capítulo 5, donde se definieron los valores de dichos parámetros en base a las pruebas realizadas.

La fase de entrenamiento genera como resultado un modelo SVM entrenado el cual puede ser usado para estimar la edad de nuevas imágenes faciales tomando como entrada las características provistas por la etapa de extracción y selección de características.

3.4. Diagrama funcional del mecanismo propuesto de estimación de la edad

En resumen, la visión general del mecanismo para estimación de la edad que contempla las etapas descritas en esta sección se integra en la Figura 24.

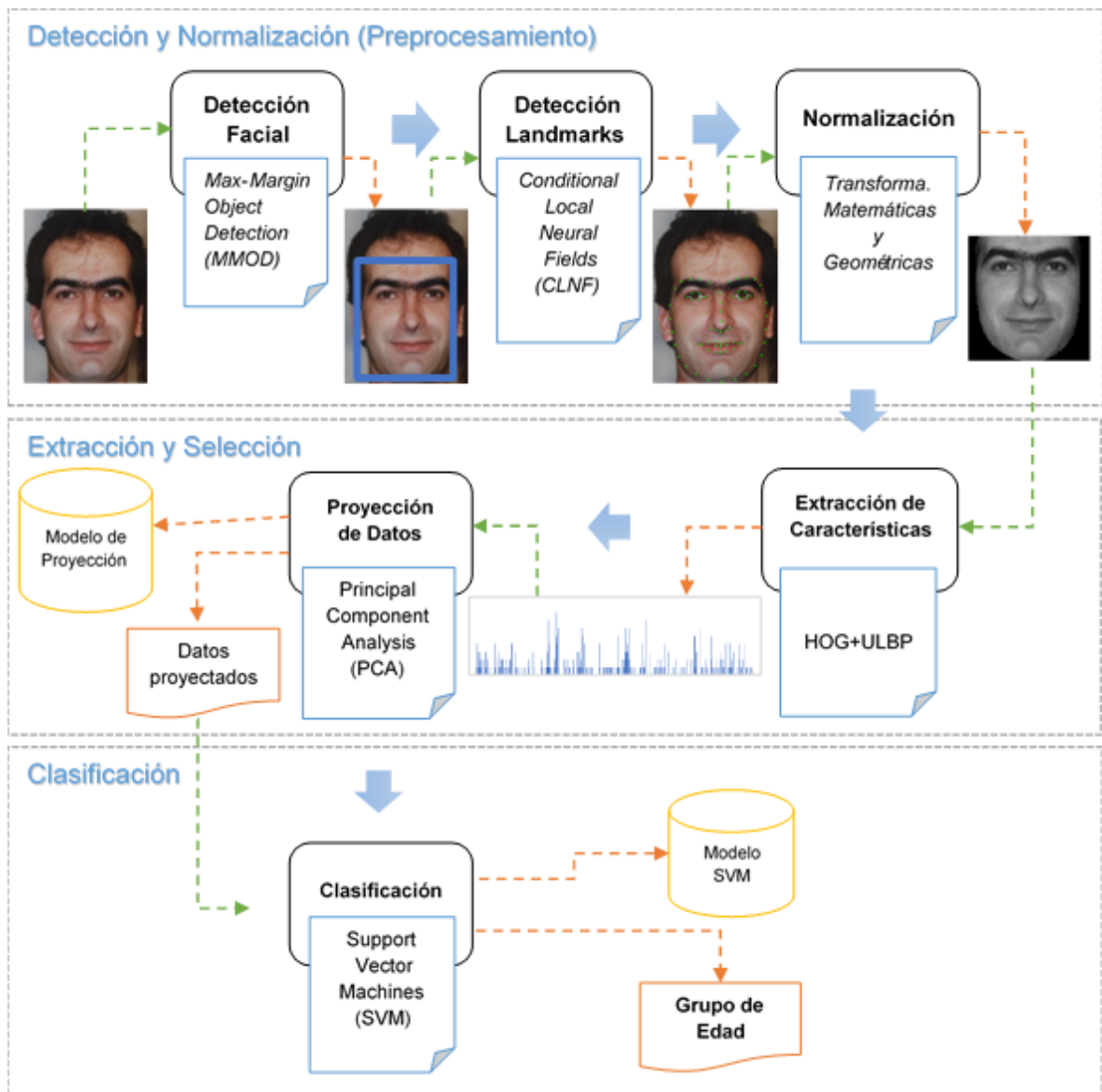


Figura 24. Diseño del mecanismo de estimación de edad

En la Figura 24, se presenta una vista completa del diseño del mecanismo de estimación de edad propuesto mostrando como a partir de una imagen facial puede ser obtenida una estimación del grupo de edad al cual pertenece la persona retratada en la imagen de entrada, este resultado es el producto de la ejecución de cada etapa: detección y normalización, extracción y selección de características y clasificación. Cabe mencionar que los modelos entrenados solo son salidas de la fase de entrenamiento y en dicha fase el mecanismo no obtiene un grupo de edad como resultado, por el contrario, solo los modelos entrenados. En la fase de prueba la salida del mecanismo es el grupo de edad y en cuanto a los modelos entrenados, solo se hace uso de ellos para realizar la estimación.

4. EVALUACIÓN DEL MECANISMO DE ESTIMACION DE LA EDAD PROPUESTO

4.1. Protocolo de Evaluación Comparativa

Como fue propuesto en los objetivos del presente trabajo, la evaluación del mecanismo propuesto fue enmarcada bajo el protocolo de evaluación comparativa LOPO [81]. No obstante, las directrices de este protocolo están basadas en las bases de datos FG-Net [128] y MORPH [129] y dirigidas a la estimación de edad exacta, lo que representó un problema debido a que el desarrollo de este trabajo está enfocado en la estimación de edad en grupos o rangos de edad. Por esta razón, no es posible aplicar LOPO tal como se describe en su definición.

En este contexto, *Benckmarking Facial Image Analysis Technologies* (BeFIT), encargado de estandarizar la evaluación de las tecnologías de análisis de imágenes faciales, presenta a partir del protocolo de evaluación anteriormente mencionado, una propuesta de protocolo de evaluación comparativa para la estimación de edad en grupos [84], usando una base de datos denominada *Image of Groups* (IoG)[86], siendo entonces determinado que este será el protocolo a seguir para la evaluación de este mecanismo dadas las condiciones del mismo, en cuanto al enfoque de estimación de la edad en grupos que satisface las características del desarrollo de este proyecto.

4.1.1. Base de datos

La base de datos IoG está compuesta por 28231 caras distribuidas en 5080 imágenes grupales tomadas de Flickr. El 86% de las caras fueron detectadas mediante un detector facial, y las demás fueron etiquetadas manualmente. Cada cara contiene metadatos entre los cuales se encuentra la etiqueta del grupo de edad en el que ha sido clasificada y el género. Las caras son clasificadas en una de siete categorías posibles: 0-2, 3-7, 8-12, 13-19, 20-36, 37-65 y 66 + estos grupos son una aproximación a las diferentes etapas de la vida de un humano. En promedio las caras tienen sólo 18,5 píxeles entre los centros del ojo, y el 25% de las caras tienen menos de 12,5 píxeles. En la Figura 25, son presentados algunos ejemplos de imágenes faciales extraídas de la base de datos IoG.



Figura 25. Ejemplo de imágenes de la base de datos loG

Algunas imágenes tienen personas sentadas, acostadas o de pie sobre superficies elevadas. Las personas suelen tener gafas oscuras, oclusiones faciales o expresiones faciales inusuales.

4.1.2. Definición del conjunto de entrenamiento y prueba

Las imágenes faciales son extraídas de las imágenes grupales de la base de datos mediante los metadatos adjuntos a ésta, de cada cara en una imagen grupal se conocen las coordenadas del rectángulo que la comprende, dichas coordenadas son el punto de referencia para extraer cada cara en una imagen diferente. Como se mencionó anteriormente, la base de datos contiene 28231 caras clasificadas en siete grupos de edad diferentes, pero no están uniformemente distribuidas dentro de los grupos, por ende, el protocolo define tomar un número fijo de imágenes por grupo seleccionadas de forma aleatoria tanto para el entrenamiento como para las pruebas. De cada grupo de edad se seleccionan 500 imágenes faciales para el entrenamiento, es decir 3500 en total, y, 150 para las pruebas, 1050 en total. De esta forma obtenemos un conjunto de datos distribuido de forma uniforme.

4.1.3. Medición de Precisión

La medición de la precisión del mecanismo de estimación se debe realizar mediante una matriz de confusión, además de que dicho método es utilizado en gran medida para evaluar sistemas de clasificación. Bajo este método se definen dos medidas de precisión, *Accuracy*

of an Exact Match (AEM)(42) y Accuracy of allowing an Error of One age category (AEO)(43), la primera medida refiere a la precisión del mecanismo de forma exacta sobre la categoría correcta a la que pertenece la imagen facial y la segunda, la precisión permitiendo una categoría de error por arriba o por debajo de la categoría correcta [87].

$$AEM = \frac{\check{N}_m}{N'} * 100\% \quad (42)$$

$$AEO = \frac{\check{N}_0}{N'} * 100\% \quad (43)$$

En las ecuaciones (42) y (43) \check{N}_m refiere al número de aciertos, N' , el número de imágenes de prueba y, \check{N}_0 , el número de aciertos contando las predicciones de una categoría por encima o por debajo de la correcta (por ejemplo, si estimó que la imagen facial pertenece a una persona de 3-7 cuando en realidad pertenece a una de 8-12, este se cuenta como un acierto),

4.2. Implementación del protocolo

Siguiendo el protocolo, de las 28231 imágenes faciales de la base de datos loG fueron seleccionadas de forma aleatoria 4550, de las cuales se tomaron 3500 para el entrenamiento, 500 por cada grupo de edad, y 1050 imágenes para pruebas, 150 por cada grupo de edad. Algunas imágenes bastante pequeñas impidieron que el detector facial utilizado en nuestro mecanismo logre realizar su función, por lo que las imágenes con un tamaño menor a 100 píxeles de alto y/o ancho fueron re-escaladas a un tamaño de 150 * 150 píxeles, este fue el único preprocesamiento realizado antes de pasar cada imagen por el mecanismo de estimación de edad propuesto.

Se extrajeron y evaluaron las características obtenidas mediante los descriptores LBP, CLBP y HOG y también las combinaciones LBP+HOG y CLBP+HOG. Cabe recordar que a las implementaciones LBP y CLBP se les aplicó ULBP con el objetivo de reducir el número de características y también al vector de características obtenido mediante cada descriptor y las combinaciones evaluadas se les realizó una reducción de dimensionalidad mediante PCA. En cuanto a la clasificación se evaluaron algunos de los algoritmos más utilizados en los antecedentes, entre ellos: *k-Nearest Neighbors* (kNN) [130], *Random Forest* (RF) [131], *Linear Discriminant Analysis* (LDA) [132] y SVM con el *kernel* RBF. Las pruebas fueron realizadas para cada descriptor con cada uno de los algoritmos de clasificación variando los parámetros de configuración para así evaluar su comportamiento.

kNN es un algoritmo de clasificación supervisada que estima la probabilidad a posteriori de que un elemento este dentro de una clase. La idea básica del algoritmo refiere a que cada nuevo caso se va a clasificar en la clase más frecuente a la que pertenecen sus k vecinos más cercanos [130]. En la Tabla 8 podemos apreciar el AEM y AEO obtenido con las

características extraídas para cada uno de los descriptores y el algoritmo de clasificación kNN asignado diferentes valores para k.

Tabla 8. AEM y AEO obtenidos para cada uno de los descriptores y diferentes valores de k usando el método kNN

Descriptor	k	AEM (%)	AEO (%)
LBP	2	32.19	65.52
CLBP	2	27.81	58.19
HOG	2	37.81	73.43
LBP+HOG	2	38.67	74.86
CLBP+HOG	2	35.43	74.86
LBP	4	37.24	71.24
CLBP	4	29.81	63.62
HOG	4	41.71	80.00
LBP+HOG	4	41.43	79.14
CLBP+HOG	4	40.86	80.47
LBP	8	37.33	73.24
CLBP	8	33.52	67.62
HOG	8	42.48	81.71
LBP+HOG	8	44.00	82.95
CLBP+HOG	8	42.19	80.48
LBP	16	39.24	74.19
CLBP	16	33.9	66.57
HOG	16	43.62	83.24
LBP+HOG	16	43.90	82.57
CLBP+HOG	16	42.76	82.95
LBP	32	39.33	74.47
CLBP	32	35.52	68.76
HOG	32	44.00	82.57
LBP+HOG	32	45.43	83.90
CLBP+HOG	32	43.52	83.33

En la Tabla 8 se destacan los resultados obtenidos con el descriptor LBP+HOG y un k = 32, los resultados obtenidos para esta configuración fueron de 45.43% y 83.90 % para AEM y AEO respectivamente. Lo que significa que del total de imágenes faciales del conjunto de prueba clasificó correctamente el 45.43% y con un error de una categoría por encima o por debajo el 83.33%.

El siguiente algoritmo evaluado de clasificación es RF, el cual es un método de aprendizaje conjunto que funciona construyendo cierta cantidad de árboles de decisión en tiempo de entrenamiento y la salida o la decisión está basada en la moda, para la clasificación, o la media, para la regresión, de la decisión tomada individualmente por cada uno de los arboles

[131]. A continuación, en la Tabla 9 se puede observar el AEM y AEO obtenido con las características extraídas con cada uno de los descriptores y el algoritmo de clasificación RF asignado diferentes valores para el parámetro l , que refiere a la cantidad de árboles que se considerarán para la clasificación.

Tabla 9. AEM y AEO obtenidos para cada uno de los descriptores y diferentes valores de l usando el método RF

Descriptor	l	AEM (%)	AEO (%)
LBP	10	18.00	46.57
CLBP	10	20.48	48.38
HOG	10	19.24	45.80
LBP+HOG	10	18.00	44.86
CLBP+HOG	10	22.19	49.71
LBP	50	26.19	53.04
CLBP	50	22.67	53.23
HOG	50	27.14	56.29
LBP+HOG	50	24.29	55.14
CLBP+HOG	50	28.48	58.57
LBP	100	31.90	61.71
CLBP	100	28.57	60.29
HOG	100	34.00	63.43
LBP+HOG	100	29.05	59.05
CLBP+HOG	100	32.57	61.52
LBP	500	41.90	73.81
CLBP	500	37.52	68.29
HOG	500	43.81	74.86
LBP+HOG	500	40.10	74.19
CLBP+HOG	500	43.71	73.71

En la Tabla 9 se destacan los resultados obtenidos con el descriptor HOG y un l con valor de 500, los resultados obtenidos para esta configuración fueron 43.81% y 74.86 % para AEM y AEO respectivamente. Lo que quiere decir que del total del total de imágenes faciales que pertenecen al conjunto de prueba clasificó correctamente el 43.71% y con un error de una categoría por encima o por debajo el 73.71 %, resultados que se encuentran por debajo de los obtenidos para KNN.

Continuando con la evaluación realizamos un proceso similar con el algoritmo LDA, su objetivo principal es encontrar una combinación lineal de los rasgos que caracterizan dos o más clases. Dicha combinación puede usarse como un clasificador lineal o reducción de dimensionalidad [132]. En la Tabla 10 podemos apreciar el AEM y AEO obtenido con las

características extraídas con cada uno de los descriptores y el algoritmo de clasificación LDA, este algoritmo no requiere de parámetros extra.

Tabla 10. AEM y AEO obtenidos para cada uno de los descriptores usando el método LDA

Descriptor	AEM (%)	AEO (%)
LBP	49.05	82.10
CLBP	47.52	80.76
HOG	51.33	84.19
LBP+HOG	50.67	83.80
CLBP+HOG	51.14	84.10

En la Tabla 10 se destacan los resultados obtenidos con el descriptor HOG. Los resultados obtenidos para esta configuración fueron 51.33% y 84.19% para AEM y AEO respectivamente. La interpretación que se le da a este resultado refiere a que del total de imágenes que pertenecen al conjunto de imágenes faciales de prueba clasificó correctamente el 51.33% y con un error de una categoría por encima o por debajo el 84.19%, estos resultados superan los obtenidos por KNN y RF.

Finalmente, en la Tabla 11 podemos apreciar el AEM y AEO obtenido para las características extraídas con cada uno de los descriptores y el algoritmo de clasificación SVM con el *kernel* RBF, así como se mencionó en el Capítulo 3. El parámetro γ fue definido como $\frac{1}{n^{\circ} \text{ Características}}$ y los valores de C son variables.

Tabla 11. AEM y AEO obtenidos para cada uno de los descriptores y diferentes valores de C usando el método SVM

Descriptor	γ	C	AEM (%)	AEO (%)
LBP	$\frac{1}{424} \approx 0.00235849$	1	40.29	71.33
CLBP	$\frac{1}{637} \approx 0.00156986$	1	18.67	49.14
HOG	$\frac{1}{828} \approx 0.00120771$	1	58.10	89.23
LBP+HOG	$\frac{1}{907} \approx 0.00110254$	1	58.00	89.24
CLBP+HOG	$\frac{1}{836} \approx 0.00119617$	1	58.29	89.93
LBP	$\frac{1}{424} \approx 0.00235849$	2	43.90	75.33
CLBP	$\frac{1}{637} \approx 0.00156986$	2	18.67	49.14

HOG	$\frac{1}{828} \approx 0.00120771$	2	59.33	90.95
LBP+HOG	$\frac{1}{907} \approx 0.00110254$	2	59.71	91.52
CLBP+HOG	$\frac{1}{836} \approx 0.00119617$	2	59.62	91.04
LBP	$\frac{1}{424} \approx 0.00235849$	4	47.81	79.62
CLBP	$\frac{1}{637} \approx 0.00156986$	4	19.33	49.52
HOG	$\frac{1}{828} \approx 0.00120771$	4	59.24	91.52
LBP+HOG	$\frac{1}{907} \approx 0.00110254$	4	60.10	91.62
CLBP+HOG	$\frac{1}{836} \approx 0.00119617$	4	59.33	91.62
LBP	$\frac{1}{424} \approx 0.00235849$	8	49.62	81.42
CLBP	$\frac{1}{637} \approx 0.00156986$	8	24.10	59.42
HOG	$\frac{1}{828} \approx 0.00120771$	8	58.00	92.00
LBP+HOG	$\frac{1}{907} \approx 0.00110254$	8	58.95	92.00
CLBP+HOG	$\frac{1}{836} \approx 0.00119617$	8	58.48	92.10
LBP	$\frac{1}{424} \approx 0.00235849$	16	50.38	81.90
CLBP	$\frac{1}{637} \approx 0.00156986$	16	29.52	65.52
HOG	$\frac{1}{828} \approx 0.00120771$	16	57.24	91.52
LBP+HOG	$\frac{1}{907} \approx 0.00110254$	16	57.24	91.33
CLBP+HOG	$\frac{1}{836} \approx 0.00119617$	16	57.24	91.61
LBP	$\frac{1}{424} \approx 0.00235849$	32	52.10	83.33
CLBP	$\frac{1}{637} \approx 0.00156986$	32	36.00	69.90
HOG	$\frac{1}{828} \approx 0.00120771$	32	56.00	90.57
LBP+HOG	$\frac{1}{907} \approx 0.00110254$	32	56.86	90.76

CLBP+HOG	$\frac{1}{836} \approx 0.00119617$	32	55.90	90.76
----------	------------------------------------	----	-------	-------

En la Tabla 11 se destacan los resultados obtenidos con el descriptor LBP+HOG y CLBP+HOG con un valor de C de 4 y 8 respectivamente. LBP+HOG obtuvo el AEM más alto con un 60.1% y CLBP+HOG el AEO más alto con 92.10%. Esto significa que para el caso de LBP+HOG el algoritmo de clasificación logró etiquetar con exactitud el 60.1% de las imágenes faciales del conjunto de prueba, y para CLBP+HOG, con un error de una categoría por encima o debajo, logró un porcentaje de acierto de 92.10%.

A partir de las tablas 10,11,12 y 13 donde fueron presentados los resultados de evaluación de los descriptores LBP, CLBP, HOG, LBP+HOG y CLBP+HOG y, los métodos kNN, RF, LDA y SVM con el *kernel* RBF, se destacan en cada una de ellas como los valores más relevantes para AEM los siguientes: 45.43% para kNN con el descriptor LPB+HOG, 43.81% para RF con HOG, 51.33% para LDA con HOG y 60.10% para SVM con LBP+HOG. De igual forma los valores destacados para AEO refieren a: 83.90% para kNN con el descriptor LPB+HOG, 74.86% para RF con HOG, 84.19% para LDA con HOG y 92.10% para SVM con CLBP+HOG.

El mejor resultado en cuanto a aciertos de clasificación (AEM) se obtuvo mediante el descriptor LPB+HOG con el algoritmo de clasificación SVM para un C=4 con un porcentaje de acierto del 60.10%, en cuanto al porcentaje de precisión obtenido al incluir como aciertos las clasificaciones erradas por una categoría arriba o debajo de la real (AEO), la precisión es del 91.62% que no se aleja mucho del 92.10% que fue el máximo valor alcanzado en la evaluación. En base a estos resultados se ha elegido LBP+HOG como descriptor para el mecanismo de estimación, teniendo en cuenta que se aplica ULBP sobre el resultado de LBP, como se describió en el Capítulo 3 para reducir la dimensionalidad del vector de características, PCA como reductor de dimensionalidad y SVM como algoritmo de clasificación. Al ser esta la combinación elegida, en la Tabla 12 podemos apreciar en detalle la matriz de confusión obtenida como resultado de la evaluación de dicha combinación.

En la matriz de confusión presentada, se puede apreciar que la configuración funciona mejor para la clasificación en los grupos 0-2 y 66+ con un AEM de 70.67 y 76.67 respectivamente. Sin embargo, su desempeño no es tan alto en rangos como 3-7 con un AEM de 50.67.

Este tipo de resultados, es decir, con valores similares han sido obtenidos en trabajos anteriores [84][89][90], por lo que podemos igualmente asumir que el bajo desempeño se debe a los cambios de textura que no son tan notables entre estos rangos de edad.

Tabla 12. Matriz de confusión de la evaluación para el descriptor LPB+HOG y el método de clasificación SVM con el *kernel* RBF y C = 4

		Estimado								AEM	AEO
		Rangos	0-2	3-7	8-12	13-19	20-36	37-65	66+		
Real	0-2	106	30	11	1	1	0	1	70.67	90.66	
	3-7	20	76	47	4	1	1	1	50.67	95.33	
	8-12	5	24	91	25	4	0	1	60.67	93.33	
	13-19	0	6	22	82	25	3	12	54.67	86.00	
	20-36	1	1	9	29	84	21	5	56.00	89.33	
	37-65	0	2	0	9	49	77	19	51.33	96.66	
	66+	0	0	2	3	4	26	115	76.67	94.00	
	Total								60.10	91.62	

En la Tabla 13 fue realizada una comparación de la configuración propuesta para la estimación de la edad y las propuestas encontradas en la revisión de antecedentes que han utilizado un protocolo similar para su evaluación, evidenciándose resultados superiores obtenidos en este proyecto.

Tabla 13. Comparación del mecanismo propuesto con los resultados más relevantes encontrados en los antecedentes.

Enfoque	AEM (%)	AEO (%)
Context [125]	32.9	64.4
LaplacianScore [126]	35.5	74.5
Appearance [125]	38.4	71.3
FisherScore [126]	42.8	83.7
Appearance+Context [125]	42.9	78.1
Gabor + Adaboost [124]	43.7	80.7
LAR [126]	44.8	84.9
RankBoost [126]	44.8	84.5
LBP + Adaboost [124]	44.9	83.0
Boosted Gabor + SVM [124]	48.4	84.4
PLO [126]	48.5	88.0
Boosted LBP + SVM [124]	50.3	81.1
CLBP_M + SVM [121]	51.7	88.7
ML LBP + SVM [130]	55.05	88.19
ML LPQ + SVM [90]	56.0	88.8
ULBP+HOG+PCA+SVM	60.10	91.62

Como es sugerido por BeFIT, fue realizada la evaluación siguiendo los lineamientos para la estimación de edad en grupos [84] usando la base de datos IoG [86]. La Tabla 13 confirma que la configuración presentada en este trabajo para la estimación de la edad en grupos supera las encontradas en los antecedentes. Por tanto, esta configuración fue utilizada en la construcción del mecanismo de estimación integrado al prototipo de sistema de control parental presentado en este trabajo.

5. PROTOTIPO DE CONTROL PARENTAL

Comúnmente los sistemas de control parental, como se menciona en los antecedentes, poseen un amplio abanico de funciones, las cuales pueden agruparse en general en dos grandes grupos: monitoreo y control. Como fue definido en los objetivos de este trabajo, dichas funcionalidades serán orientadas al control de acceso a contenidos web desde un computador personal. Además, atendiendo a la definición de control parental como una **“Característica especialmente útil para padres y responsables educativos que desean impedir que niños o adolescentes puedan acceder a páginas Web inapropiadas”** [133]. De esta forma, el mecanismo de control parental está orientado a controlar y monitorear la navegación web realizada desde un computador personal.

A continuación, se describe el diseño e implementación de un prototipo software de control parental que integra el mecanismo de estimación de edad propuesto en el capítulo anterior. El prototipo fue construido siguiendo el Proceso Unificado Ágil (Agile Unified Process - AUP), el cual es una metodología iterativa e incremental, en la que se seleccionaron tres disciplinas: el modelo, en el que la meta es entender el problema del dominio e identificar una solución viable; la implementación, donde se transforma el modelo en código ejecutable y se realizan pruebas funcionales; finalmente, la disciplina de evaluación, que planea asegurar la calidad del prototipo mediante una evaluación objetiva. Como lenguaje de modelado se empleó el Lenguaje unificado de modelado (Unified Modeling Language – UML).

Para la construcción del prototipo se empleó la Plataforma de Windows Universal (*Universal Windows Platform* - UWP), para el desarrollo del mecanismo de estimación de edad mediante imágenes faciales se utilizó el conjunto de clases de OpenCV y OpenFace. El prototipo construido permite restringir el acceso a contenido web, mediante imágenes faciales obtenidas desde la cámara del computador personal.

5.1. Requerimientos

Se identifican los requerimientos a alto nivel, se crea una lista de requerimientos y se configura el ambiente de desarrollo, siendo estos:

- Definir reglas para los grupos de edad: el usuario debe seleccionar las categorías que se restringen para cada grupo de edad de usuarios.
- Monitorear eventos de red: el usuario visualiza los eventos generados por el filtrado de la interfaz de red, según las reglas definidas.
- Estimar la edad de las imágenes faciales de un segmento de video. El sistema debe poder estimar la edad en un intervalo de tiempo, con el objetivo de establecer el

grupo al que pertenecen las reglas que deben actuar sobre el filtro de paquetes de red.

5.1.1. Otras consideraciones

El prototipo debe establecer un mecanismo mediante los privilegios por usuarios del sistema operativo, con el objetivo de evitar que se detenga el control sobre la red y, por lo tanto, la restricción a contenidos web. Para lograr esto, debe ser construido utilizando herramientas que faciliten la programación de prestaciones del sistema operativo Windows.

5.2. Análisis y Diseño

A continuación, se presentan los artefactos que permitieron establecer un modelo arquitectónico viable de la solución software del prototipo. Mediante casos de uso, se definieron los posibles escenarios donde los usuarios del sistema y el software van a interactuar entre sí.

5.2.1. Definición de actores

Se definen los actores del sistema, teniendo en cuenta su participación directa en la interacción con el software, es decir, los actores que manipulan el prototipo. Por este motivo, no se ha incluido al “usuario” del computador (quien navega y se le restringe el acceso al contenido), como un actor directo, puesto que dicha restricción de red es una consecuencia de la manipulación de paquetes de datos.

Tabla 14. Descripción de Actor Administrador

Actor ACT-01	Administrador
AUTORES	Wilian Pantoja, Alexis Ruano
DESCRIPCIÓN	El usuario Administrador podrá gestionar las reglas que se indiquen para cada grupo de edad. Este usuario visualiza la actividad web (Eventos) generados por el Controlador de red. Puede consultar el estado del servicio en segundo plano. Se le permite ver información de configuración del control parental.
COMENTARIOS	

Tabla 15. Descripción de Actor Controlador de Red

Actor ACT-02	Controlador de red (Sistema)
---------------------	------------------------------

AUTORES	Wiliam Pantoja, Alexis Ruano
DESCRIPCIÓN	El sistema toma el control sobre la interfaz de red del computador y filtra los paquetes según las reglas establecidas por el Actor Administrador. El sistema genera eventos que describen el paquete interceptado y el contenido encontrado y que coincide con las reglas por cada grupo de edad.
COMENTARIOS	El controlador de red intercepta los paquetes de las peticiones web de un usuario del computador (mediante un navegador de internet, por ejemplo). Por tanto, el usuario es un actor indirecto del sistema que se ve afectado por las decisiones que el sistema tome respecto al acceso al contenido web.

5.2.2. Casos de uso

Los casos de uso están modelados bajo dos escenarios principales: El escenario de la configuración y monitoreo del sistema de control parental se muestra en la Figura 26 y el escenario de filtrado de paquetes en la Figura 27. Los casos de uso en formato extendido se pueden apreciar con detalle en el **Anexo A**.

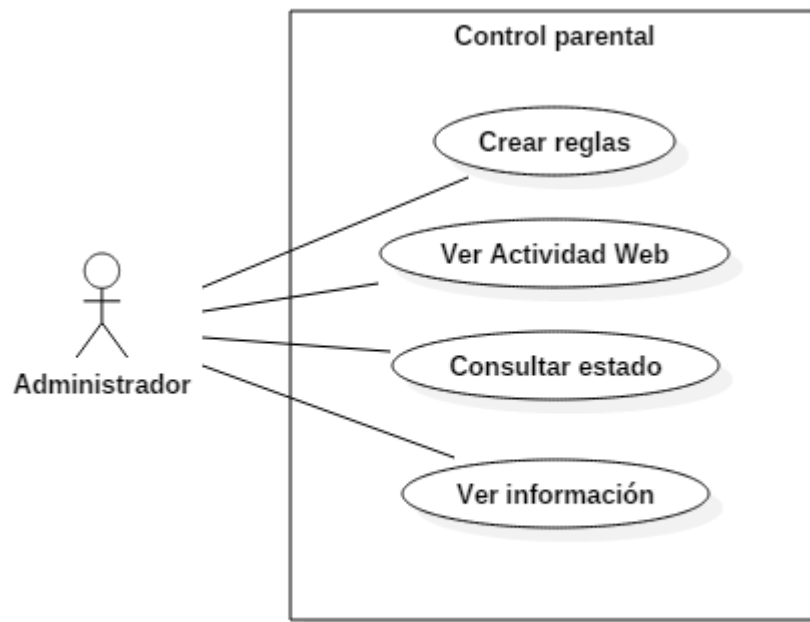


Figura 26. Diagrama de casos de uso para el escenario de configuración y monitoreo

En la Figura 26 podemos ver los casos de uso para el Actor Administrador. Los casos de uso comprenden la creación de reglas, ver el registro de la actividad web hecha en el ordenados, consultar el estado del sistema de control parenta y ver la información del sistema. Este escenario comprende las posibilidades que ofrece el sistema mediante la interface de usuario.

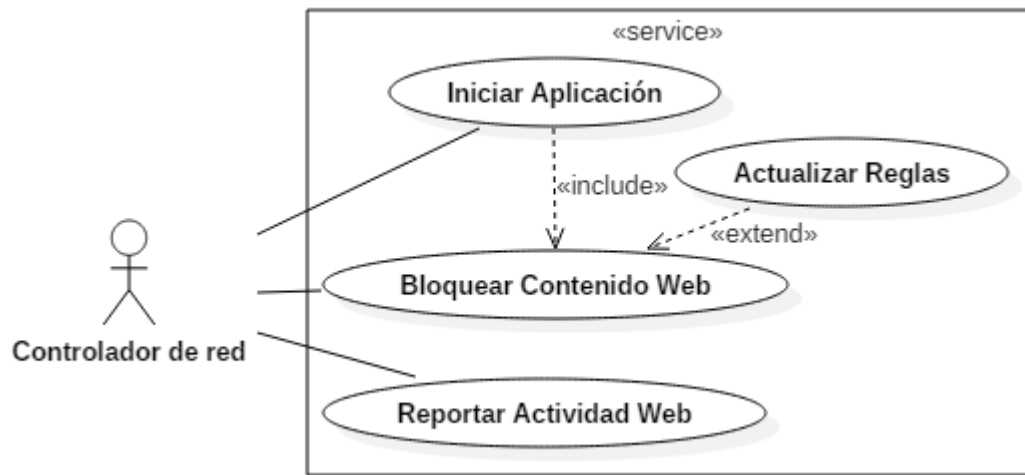


Figura 27. Diagrama de casos de uso para escenario de servicio de filtrado de paquetes de red

En la Figura 27 se describen los casos de uso para el controlador de red el cual se va a encargar de invocar el inicio de la aplicación, bloqueo de contenido web y hacer el registro de actividad web. Recordemos que el actor denominado como controlador de red no es un ente humano, sino un proceso del sistema operativo.

5.2.3. Arquitectura del sistema

Una vez se definieron los casos de uso de alto nivel y requerimientos funcionales se procede con la configuración del ambiente de desarrollo y el diseño de la arquitectura a alto nivel.

A continuación se expone la guía arquitectónica empleada para la construcción del prototipo a través del conjunto de vistas planteado en el Modelo de vistas 4+1 de RUP (*Rational Unified Process*) [21]. Para la definición del conjunto de vistas (vista de casos de uso, lógica, implementación, procesos y despliegue) fueron tenidos en cuenta los artefactos de las disciplinas de modelo, implementación y pruebas más relevantes.

5.2.3.1. Vista de casos de uso

Esta vista representa la manera como los usuarios interactúan con la aplicación. A continuación, se describen los escenarios que representan las principales funcionalidades del prototipo.

5.2.3.1.1. Crear Reglas

Esta funcionalidad permite asociar los grupos de edad con las categorías de contenido web que el sistema permite interceptar. Las reglas son almacenadas en una base de datos local y son el punto de partida para la generación de eventos, puesto que se generan las listas de acceso para que el controlador de red rechace o permita el tránsito de paquetes de red.

5.2.3.1.2. Bloquear contenido web

El controlador de red mediante el componente de estimación de la edad determina la edad de la persona que está navegando en el computador y obtiene las listas de acceso según la edad estimada. Según estas listas de acceso la librería de filtrado de red rechaza los paquetes de red, considerándolos como eventos que son registrados en la base de datos local.

5.2.3.1.3. Ver actividad web

Los eventos generados por el controlador de red son listados al actor Administrador y se le permite ver información complementaria como la regla asociada al evento, la información de ocurrencia temporal, e información relevante del paquete interceptado.

5.2.3.2. Vista lógica

Ésta vista considera los componentes arquitectónicos de la solución software, como capas, módulos y paquetes de arquitectura.

5.2.3.2.1. Componentes arquitectónicos

La arquitectura del prototipo de Control Parental construido se presenta en la Figura 28. El componente de Interfaz gráfica de usuario (*Graphical User Interface* - GUI) mediante *Universal Windows Platform* (UWP) representa la interfaz gráfica del Control parental para el administrador, desarrollada en la plataforma universal de Windows. Por medio de ésta el administrador podrá visualizar la información requerida y asignar reglas según los grupos de edad. Los componentes de estimación de edad y controlador de red son los encargados del núcleo del sistema ya que son los usados por el servicio de control parental para realizar su función. En el componente de estimación se encuentra toda la lógica que permite realizar

desde la detección facial hasta la clasificación de la imagen facial dentro de uno de los grupos o rangos de edad establecidos. Mediante el controlador de red se realiza la lectura de paquetes de red, lo cual permite acceder a la información necesaria para procesar el tránsito de red mientras un usuario navega normalmente por internet. Los componentes de sistema operativo y hardware se refieren a los recursos requeridos tanto de memoria compartida como de dispositivos físicos (cámara web e interfaz de red).

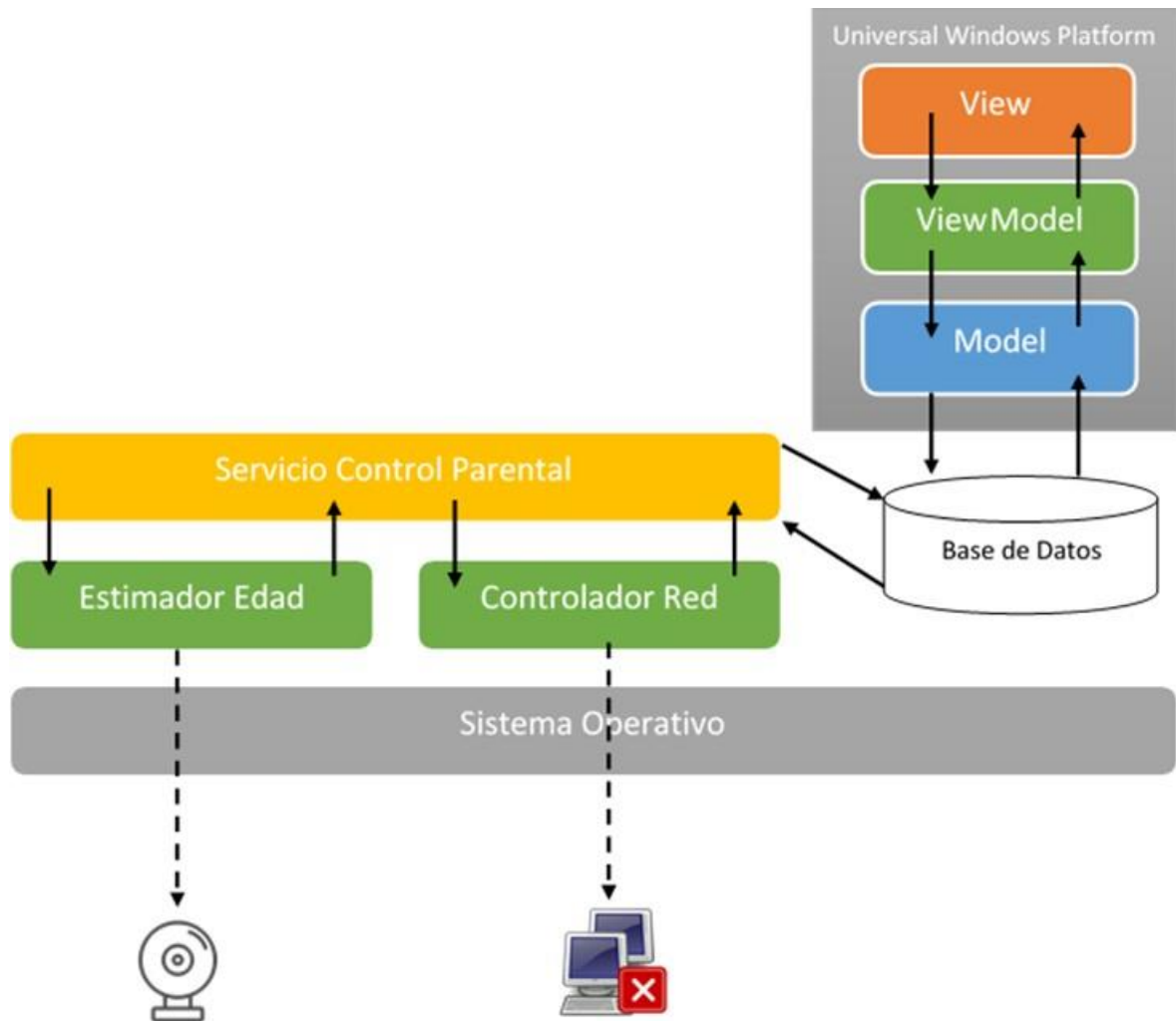


Figura 28. Arquitectura del Prototipo

El componente de base de datos permite persistir los eventos generados por el controlador de red y permitir ser visualizados en el control parental, además almacenar las categorías de contenido web y las listas de acceso, dicha información es presentada y manipulada por el administrador mediante el componente desarrollado en UWP.

5.3. Implementación del prototipo

El desarrollo del prototipo se describe a continuación y se muestran los productos y actividades logradas a través del ciclo de diseño y construcción.

5.3.1. Vista de implementación

Aquí se describen los componentes que conforman la arquitectura descrita en el diseño de alto nivel.

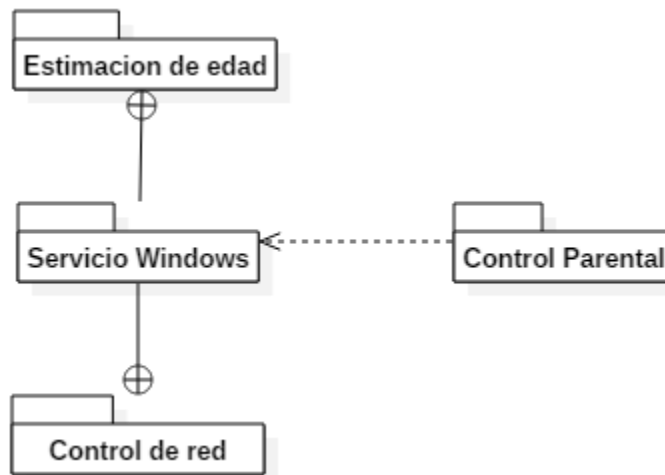


Figura 29. Diagrama de paquetes de software

En la Figura 29 se muestra el diagrama de paquetes de software. El control parental contiene los componentes de interfaz gráfica y de acceso a la base de datos. El paquete de Servicio Windows contiene las instrucciones necesarias para ejecutar el servicio en el sistema operativo y dejar ejecutando en segundo plano el control de la red y la estimación de la edad mediante imágenes faciales. Los paquetes de software de control de red y estimación de edad están contenidos dentro del Servicio Windows, mientras que control parental depende del Servicio para visualizar los eventos y las reglas especificadas.

5.3.1.1. Librería de estimación de edad

Este componente de software se crea como una librería de clases y funciones escrita en lenguaje C++. Este lenguaje es utilizado por OpenCV para implementar sus funcionalidades de operaciones de matrices y reconocimiento de patrones. El proceso que se implementa en este componente se detalla de forma teórica en el capítulo de diseño de mecanismo de estimación de edad.

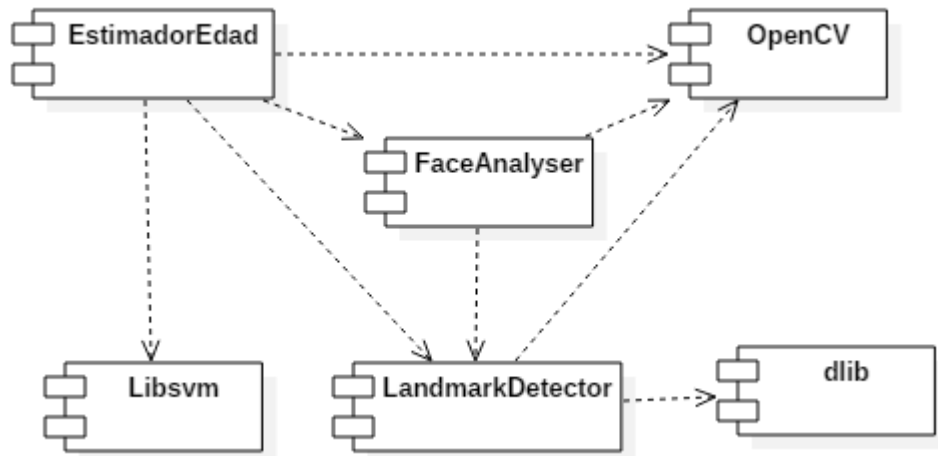


Figura 30. Diagrama de componentes del componente de estimación de edad

Los componentes de la Figura 30, hacen parte del paquete de software de estimación de edad. El componente EstimadorEdad depende de los componentes FaceAnalyser, OpenCV y Libsvm debido a que contienen los algoritmos de procesamiento de imágenes y de clasificación necesarios para el cálculo de la edad. El componente LandmarkDetector detecta el rostro sobre una imagen facial y a su vez depende de dlib (librería con algoritmos de aprendizaje de máquina).

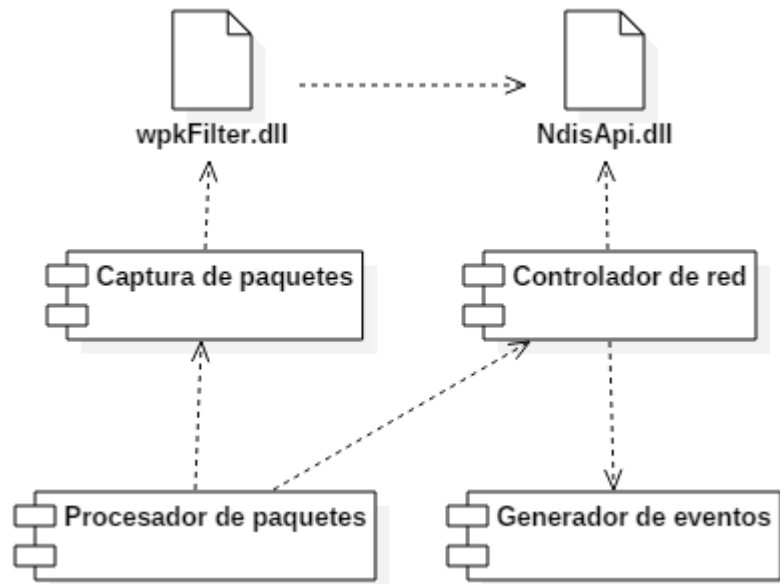


Figura 31. Diagrama de componentes del control de red

Los componentes de la Figura 31, son descritos a continuación:

- Captura de paquetes: Este componente utiliza una librería que permite el monitoreo y control sobre los paquetes de red para recibir y preparar los datos extraídos del paquete de red.
- WpkFilter.dll: es una librería de terceros, de filtrado de paquetes de red, implementada para desarrolladores que requieran el monitoreo de la red y la intervención en los paquetes de red.
- Procesador de paquetes: Se encarga de leer los paquetes y obtener su contenido, para buscarlo dentro de la lista de control de acceso. En caso de encontrarse coincidencia de un elemento de la lista de acceso con el contenido del paquete (dirección URL solicitada por el usuario mezclada con otros datos), se procede al bloqueo del paquete.
- Controlador de red: Este componente recibe el paquete que debe ser bloqueado y guarda algunos datos relevantes en una estructura temporal. Se construye un nuevo paquete de red y se envía en reemplazo del anterior.
- Generador de Eventos: De manera asíncrona se envía una estructura temporal con información del paquete y la regla asociada, para ser almacenado localmente en la base de datos.

El paquete de control de red contiene el componente controlador de la red y es el encargado de preparar los paquetes capturados, luego de procesarlos mediante una búsqueda de la información del paquete dentro de los dominios contenidos en las categorías (reglas), mediante una búsqueda paralelizada (aprovechando recursos del computador con hilos) se localizan las incidencias de las reglas, y finalmente entregarlos al componente generador de eventos para almacenarlo en la base de datos.

5.3.1.2. Servicio del sistema operativo

Es el encargado de registrar la ejecución en segundo plano del monitoreo de red y captura de imágenes. Este software se ejecuta una sola vez en el sistema operativo del computador personal, y registra el servicio para que inicie de forma automática (al inicio del sistema operativo).

En este paquete de software el usuario Administrador instala en forma de servicio en el sistema operativo el controlador de red y estimador de edad. El servicio mantiene un programa en segundo plano leyendo los paquetes de red de la conexión establecida a internet. Al ser instalado por el administrador del sistema, el resto de usuarios del computador personal no están habilitados para cerrar o desinstalar el control parental.

5.3.1.3. Base de datos de listas de acceso: Fuentes de datos

Las direcciones URL, dominios web y categorías a las que pertenecen se usan para controlar el acceso al contenido de acuerdo al grupo de edad.

Una lista de acceso es un mecanismo de control de acceso básico que permite pasar todas las URLs excepto aquellas explícitamente mencionadas [134].

La Gestión de categorías y excepciones: De forma general el usuario puede administrar los contenidos que esta permitidos visualizarse por medio del control parental, además de listas de sitios web permitidos (excepciones):

- *Blacklist*: Sitios web no permitidos (bloqueados).
- *Whitelist*: Sitios web permitidos incluso si existen en la base de datos de control de acceso.

La base de datos de control de acceso está basada en una de las listas negras que el sitio web SquidGuard [135] usa para el filtrado en el software proxy “Squid”.

5.3.2. Proceso de desarrollo

Desde el inicio de este proyecto, se planteó una solución ejecutable en el sistema operativo Windows, debido a que es bien conocido que Windows abarca gran parte del mercado de computadores personales, por tanto, se utilizan librerías, frameworks y entornos de desarrollo especializados para crear aplicaciones en la plataforma de Windows 10. Para la implementación de todos los paquetes de software se empleó el entorno de desarrollo Visual Studio 2017, un entorno de desarrollo que facilita en gran medida la creación de interfaces gráficas y cuenta con funcionalidades de inspección de código en tiempo de ejecución (debug). El framework de procesamiento de imágenes es OpenCV, ya que cuenta con una excelente documentación y una amplia comunidad de desarrollo, además de ser una librería de software abierto y libre.

Para la construcción de este prototipo se realizaron 4 iteraciones. En la primera iteración se convierten los archivos que contienen las categorías, dominios y direcciones URL de las listas de bloqueo a tablas en SQL y se crea un archivo script SQL, para crear la base de datos SQLite. Por otro lado, se prueban algoritmos de clasificación y de detección facial.

En la segunda iteración se crean las vistas del control parental y se inicia con la programación de la selección de categorías por grupos de edad, también se descargan algunas bases de datos faciales para entrenamiento y prueba de los algoritmos de clasificación.

Posteriormente en la tercera iteración, se crea el servicio Windows que integra la librería de que contiene el mecanismo de estimación de edad y el controlador de red. En el prototipo

se desarrolla el monitor de eventos generados con base en las reglas por cada grupo de edad. Se probaron algoritmos de clasificación como KNN, RF, SVM, y se decide emplear SVM para clasificación

En la cuarta iteración se corrigen errores detectados en la interfaz gráfica y el diseño de la base de datos, y se crea el servicio que integra los componentes software del mecanismo de estimación de edad y controlador de red.

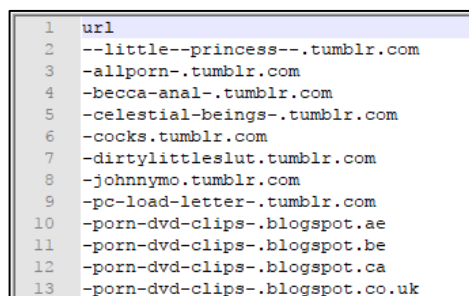
Adicionalmente, en el lenguaje de programación C++ fueron implementados los paquetes de software de servicio de Windows, Estimador de edad y controlador de red. En el lenguaje C# se implementó el control parental. La base de datos (con base en unos archivos de texto) esta implementada con el gestor de base de datos SQLite.

5.3.2.1. Primera iteración

En esta iteración se configuró el entorno de desarrollo, se instalaron las librerías de procesamiento de imágenes y la librería de filtrado de paquetes de red realizó la búsqueda de los métodos de estimación que han sido tendencia para cada una de las etapas del proceso de estimación de la edad, desde la detección facial hasta la clasificación.

5.3.2.1.1. Exploración de listas de acceso

Se investiga sobre bases de datos que agrupen URLs en categorías. SquidGuard es un redirector de URL que utiliza *blacklists* (listas negras) con el software proxy Squid. En esta etapa la lista de bloqueo encontrada se carga en una base de datos que es utilizada para hacer las búsquedas de las peticiones que son interceptadas por una librería que sea capaz de leer los paquetes que transitan sobre la interfaz de red del computador persona.



```
1 url
2 --little--princess--.tumblr.com
3 -allporn-.tumblr.com
4 -becca-anal-.tumblr.com
5 -celestial-beings-.tumblr.com
6 -cocks.tumblr.com
7 -dirtylittleslut.tumblr.com
8 -johnnymo.tumblr.com
9 -pc-load-letter-.tumblr.com
10 -porn-dvd-clips-.blogspot.ae
11 -porn-dvd-clips-.blogspot.be
12 -porn-dvd-clips-.blogspot.ca
13 -porn-dvd-clips-.blogspot.co.uk
```

Figura 32. Ejemplo de un archivo contenedor de las URL que pertenecen a una categoría

Seguidamente de encontrar la lista de acceso adecuada, se transforma en forma de base de datos, puesto que la conformación actual es en archivos de texto, en la Figura 32 se detalla una muestra del archivo de dominios en la categoría “adult”.

El texto plano que se guarda en memoria es almacenado en una base de datos que facilita búsquedas en grandes cantidades de datos. El motor de base de datos escogido es SQLITE por su facilidad de trabajar en sistemas embebidos, dispositivos móviles y por ser un motor liviano, puesto que en el caso de este proyecto la lógica relacional no necesita de características críticas.

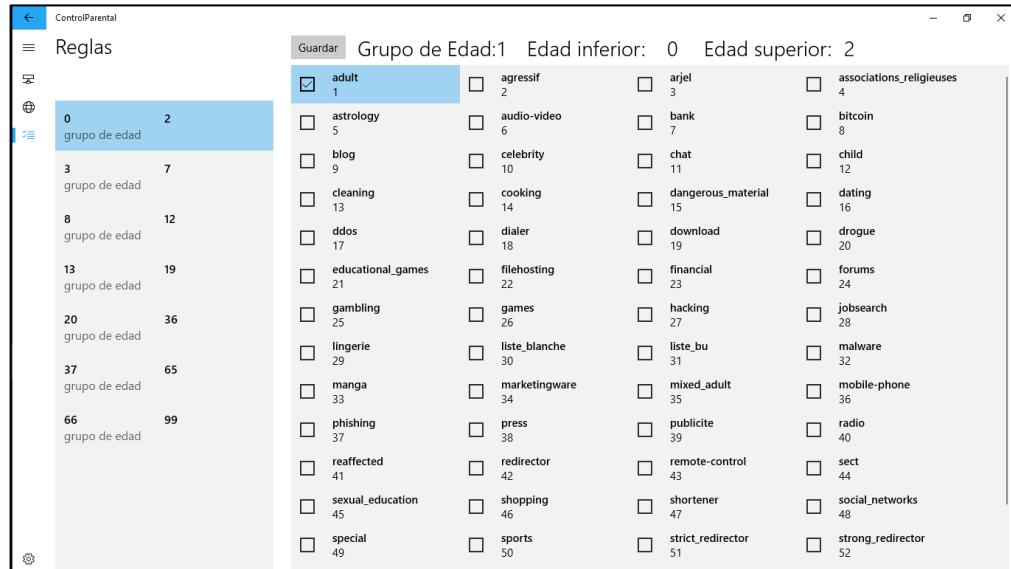


Figura 33. Configuración de reglas para un grupo de edad.

En la Figura 33, se muestran algunas de las categorías que fueron convertidas desde los archivos descargados de SquidGuard, y transformados en formato relacional y almacenado en una base de datos local.

5.3.2.1.2. Exploración de librerías de red

Se inició con la búsqueda y selección de la librería que permita la manipulación de paquetes y monitoreo de la interfaz de red. La librería *Windows Packet Filter* © de NTKernel, es un framework de filtrado de paquetes con alto rendimiento para el sistema operativo Windows®, que permite de filtrar (monitoreo y modificación) de forma transparente los paquetes de red en bruto al nivel NDIS de la capa de red con un mínimo impacto en la actividad de la red y sin tener que escribir un código de driver a bajo nivel de programación. Los requerimientos técnicos se pueden detallar en la página web de *Windows Packet Filter* [136]. La principal ventaja de esta librería es la habilidad para la manipulación de frames Ethernet en bruto, logrado mediante la instalación de un driver por debajo de todos los protocolos de red, y por encima del driver de interfaz de red. De esta forma el driver tiene el control sobre los paquetes de entrada y de salida (flujo del tráfico de red), permitiendo modificar cualquier paquete, omitirlo (bloquearlo), e incluso olvidarlo e insertar uno completamente nuevo.

5.3.2.2. Segunda Iteración:

Una vez realizado el proceso de investigación y configuración del entorno para la implementación del mecanismo de estimación de edad, se procedió a implementar el proceso de extracción de características con la ayuda de las librerías OpenFace [117] y OpenCV [28]. Los detalles de dicha implementación fueron expuestos en el Capítulo 3.

En la Figura 34 se pueden ver algunas imágenes tomadas respecto a las pruebas de escritorio del proceso de extracción de características. Como resultado del proceso para la etapa de entrenamiento, tenemos un archivo con extensión csv, el cual describe todas las características extraídas con los descriptores LBP aplicando ULBP y HOG. Dicho archivo de entrenamiento es manipulado para que sea posible usarlo desde RapidMiner [137] como entrada para hacer el proceso de evaluación respectivo.

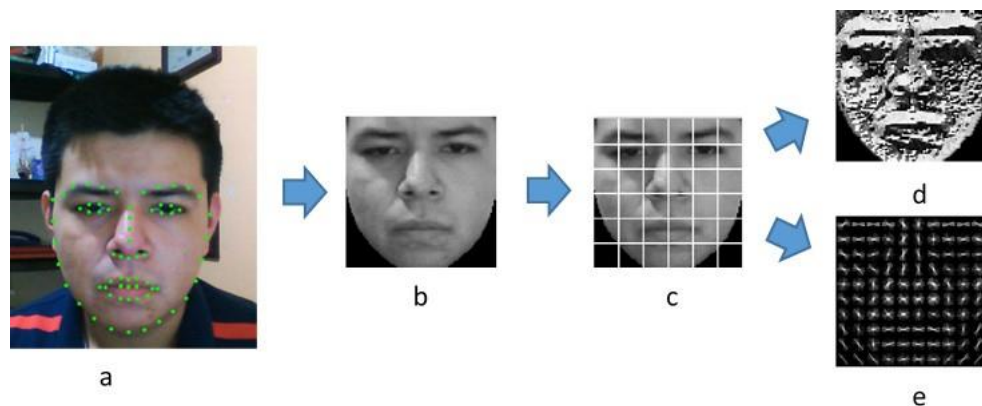


Figura 34. Ejemplo del proceso de extracción de características: a) Detección de cara y landmarks. b) Escala de grises y recorte. c) Separación en celdas. d) Aplicación del operador LBP. e) Vectores de gradiente para cálculo de HOG

Los procesos de evaluación diseñados en RapidMiner pueden observarse en las Figuras 17,18, 19 y 20 del Anexo D, donde cada modelo consta de 2 entradas etiquetadas como *Retrieve* y *Retrieve Test*, las cuales corresponden a los datos de entrenamiento y prueba respectivamente. El proceso inicia con la lectura de los datos provenientes de *Retrieve* (entrenamiento) los cuales son utilizados como datos de entrada para el aprendizaje del respectivo algoritmo de clasificación, kNN, RF, LDA o SVM según sea el caso, dicho algoritmo genera un modelo el cual es insumo para la siguiente etapa, *Apply Model*, en la cual se realiza la clasificación de los datos de prueba provenientes de *Retrieve Test* (prueba). Finalmente, en la etapa de *Performance* se evalúan los resultados arrojados por la etapa *Apply Model* y como salida de dicha etapa se obtiene una matriz de confusión la cual expone los resultados de la clasificación en cuanto a precisión.

Estos procesos de evaluación presentan como resultado una matriz de confusión, la cual es interpretada y comparada con todos los resultados como se describe en el Capítulo 4.

Según estos resultados se decide implementar SVM con el *kernel* RBF y un $C=4$ como clasificador para el mecanismo de estimación de edad. Dicha implementación se realiza con ayuda de la librería Libsvm [125].

5.3.2.2.1. Gestión de eventos de alerta en el módulo de control de red

Los eventos generados por el control de red se almacenan en la base de datos local. Cada evento contiene la información de la regla que coincide con el paquete de red detectado, la fecha y la edad detectada.

5.3.2.2.2. Paralelismo en proceso de búsqueda de URL (*Request*) en La lista negra

De acuerdo con el grupo de edad estimado para el usuario, se obtiene a partir de la base de datos local, una lista con la lista negra de dominios que deben ser comparados con la información del paquete de red. La lista se almacena en memoria en un contenedor de la librería estándar de C++, de manera que para realizar la búsqueda en esa estructura se debe hacer de forma lineal, por lo que la complejidad computacional de esta operación es de una cota del valor de elementos del vector $O(n)$. Este problema se soluciona mediante la implementación de una búsqueda con la ayuda de la librería PPL (*parallel pattern library*) de Windows, que permite el procesamiento paralelo a un alto nivel de programación.



Figura 35. Página web que es mostrada cuando se detecta el acceso a una URL del control de acceso (lista negra).

El flujo de la red es redireccionado a una página con un mensaje personalizado. La Figura 35, muestra el resultado visible al usuario del computador personal, cuando intenta acceder a contenido restringido de acuerdo con su edad.

5.3.2.3. Tercera Iteración

En esta iteración se implementa la clasificación para la solución final, en el evaluó la configuración propuesta, y teniendo claro cuales métodos iban a conformar el mecanismo, se procedió a hacer la implementación de la etapa de clasificación para la solución final.

```
resultados:array = [];  
modelos:array = [];  
si(imagen.vacio() == falso){  
    imagenEscalaGris:TipoImagen = convertirEscalaGris(es:Imagen);
```

```

deteccionFacial:booleano = detectarLandMarks(imagenEscalaGrises);
si(deteccionFacial == verdadero){
    modelos = establecerModelosActivos();
    para_cada(modelo en modelos):
        imagenRecortada:tipoImagen = recortarImagen(imagenEscalaGrises);
        celdas:array = dividirEnCeldas(imagenRecortada);
        caractLBP:array = extraerCaracteristicasLBP(celdas);
        caractHOG:array = extraerCaracteristicasHOG(imagenRecortada);
        caractLBPNormalizadas:array = normalizar(caractLBP,0,1);
        caractHOGNormalizadas:array = normalizar(caractHOG,0,1);
        caracteristicas:array = concatenar(caractLBPNormalizadas,
caractHOGNormalizadas);
        caractPCA:array = modeloPCA.proyectar(caracteristicas);
        caractPCAEscaladas:array = modeloEsala.escalar(caractPCA);
        grupoEdad:Entero = modeloSVM.predecir(caractPCAEscaladas);
        resultados.agregar(grupoEdad);
    fin para_cada;
}
}
retornar resultados;

```

Algoritmo 1. Estimación de la edad a partir de una imagen facial.

En el Algoritmo 1, se muestra el algoritmo para calcular la estimación de la edad a partir de una imagen de entrada.

5.3.2.4. Cuarta Iteración

En esta iteración se completa la integración del mecanismo de estimación de edad, representado a nivel de implementación como una librería y, el sistema de control parental que fue concebido en dos componentes principales, el componente de captura de paquetes de red o controlador de red y el componente desarrollado en UWP los cuales podemos apreciar en la Figura 28.

5.3.2.4.1. Integración de mecanismo de estimación de edad al componente de captura de paquetes de red e imágenes faciales

El mecanismo de estimación de edad es capaz de estimar el grupo de edad al que pertenece un rostro detectado en una imagen facial, sin embargo, la estimación debe hacerse en tiempo real. Por tanto, se implementa una funcionalidad que estima el grupo de edad de un grupo de imágenes, es decir, de un segmento de video del *stream* capturado de la cámara del dispositivo. Esta funcionalidad se describe a continuación:

- Se establece un periodo de 10 segundos de captura de imágenes. En este periodo de tiempo se almacenan las imágenes en memoria.
- Cada imagen es procesada con la librería de estimación y se obtiene el resultado de estimación.
- Todos los resultados de las imágenes son agrupados y ordenados de forma ascendente.
- Se calcula la moda (valor con mayor frecuencia) del arreglo de resultados de las estimaciones por cada imagen del segmento de video.
- La moda representa la estimación del grupo de edad de las imágenes faciales.
- Se obtienen las reglas asociadas a ese grupo de edad estimado (la lista de acceso).
- Los paquetes capturados son procesados comparándolos con la lista de acceso según las reglas asociadas.
- Se generan los eventos según las reglas asociadas y se almacenan en la base de datos.

5.3.2.4.2. Implementación de instalador de servicio Windows

La integración del control de red con el mecanismo de estimación de edad debe ser ejecutado en segundo plano mientras la cámara del dispositivo está encendida. Por esto, se crea una aplicación en C++ que registra un servicio llamado “ControlParental” en el sistema operativo.

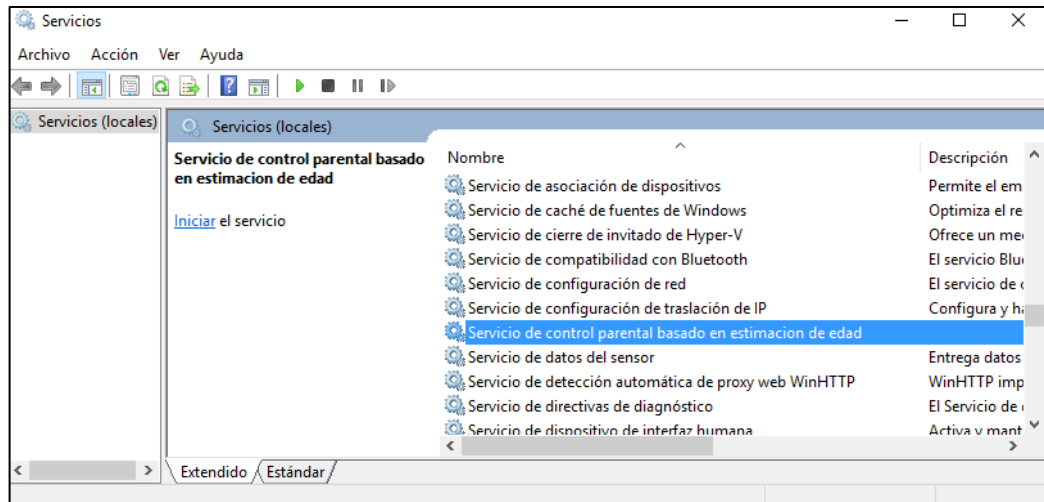


Figura 36. Servicio de prototipo de control parental instalado en la estación de trabajo.

En la Figura 36, puede observarse el servicio de control parental instalado en una estación física. El servicio se instala por el Administrador del sistema operativo, agregándose así seguridad al código en ejecución, de manera que solo podrá desinstalar o detener el

servicio, quien tenga credenciales de Administrador del sistema operativo, por tanto, un usuario sin privilegios de administración no podrá evadir el control parental.

5.3.2.4.3. Pruebas de software

Durante la última iteración se aplicaron las pruebas funcionales descritas en detalle en el **Anexo B**. Dichas pruebas fueron diseñadas para evaluar los escenarios principales de cada caso de uso del prototipo y de esa forma dar un diagnóstico general del mismo. Cada caso de prueba se compone de:

Descripción: Un breve enunciado acerca de los objetivos y desarrollo del caso de prueba.

Condiciones de ejecución: Descripción de las condiciones de ejecución que se deben cumplir antes de iniciar el caso de prueba.

Entrada: Descripción paso a paso de la ejecución del caso de prueba.

Resultado esperado: Descripción del resultado que se espera obtener al realizar el caso de prueba.

Evaluación de la Prueba: Estado del caso de prueba, el cual puede tomar algunos valores como: propuesta, pendiente de evaluación, realizada y satisfactoria, entre otros.

Se realizó la aplicación de los casos de prueba y, durante la ejecución del caso de prueba para el caso de uso bloquear contenido web, se encontró un error en la implementación. Dicho error consiste en que después de un corto periodo de tiempo (1 minuto aprox.) de iniciada la aplicación, esta deja de funcionar. La fuente de este error refiere a la conexión con la base de datos SQLite y es originado por tratar de realizar conexiones múltiples a la misma. La solución que se implementó asegura una sola conexión a la base de datos y fue basada en el patrón de diseño *singleton* y *command* para encolar las consultas a la base de datos (Figura 41).

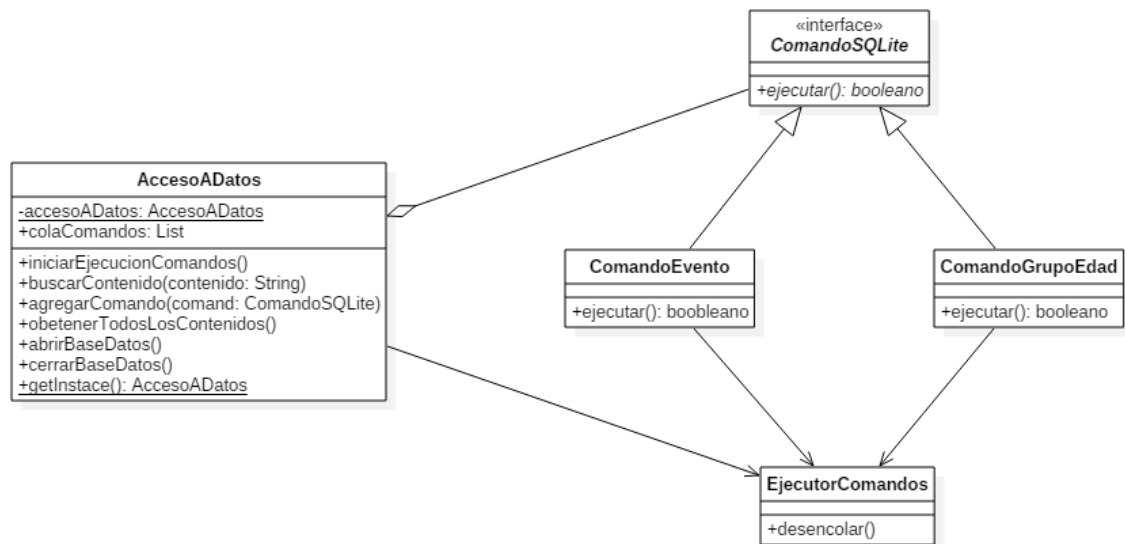


Figura 37. Mecanismo de sincronización para la ejecución de comandos sobre la base de datos.

De esta forma una instancia de la clase *AccesoADatos* se encarga de realizar la comunicación entre el programa y la base de datos (*singleton*), y a su vez se es responsable de encolar las peticiones a la base de datos, representadas bajo la interfaz *ComandoSQLite* y ejecutadas a través de la clase *EjecutorComandos* (*command*).

En cuanto a los casos de prueba restantes, la mayoría se encuentran en estado satisfactorio, a excepción de los ítems correspondientes al caso de prueba consultar estado, ya que no cumplieron con el resultado esperado. Las funcionalidades que no fueron aprobadas se encuentran en revisión.

Además de la aplicación de los casos de prueba sobre el prototipo, se realizaron pruebas de escritorio del prototipo bajo ambientes no controlados con personas con siete (7) personas con edades entre los 6 y 86 años elegidas aleatoriamente, quienes, durante la ejecución del prototipo, fueron ubicadas frente a la cámara del ordenador de forma individual. Estas pruebas permitieron identificar que la restricción de las páginas web es realizada correctamente según las reglas establecidas para cada edad y, en cuanto a la estimación de la edad, la misma depende de factores como pose y distancia de la cámara, lo cual conllevó al desarrollo del proceso descrito en la sección 5.3.2.4.1 (Integración de mecanismo de estimación de edad al componente de captura de paquetes de red e imágenes faciales). A partir de dichas pruebas, se identificó igualmente que el mecanismo no prevé falsificaciones (falsos positivos) lo cual no es erróneo, dado que la identificación de imágenes que provienen de una fotografía física o la presencia de máscaras es un tema de investigación propio, adicional al objeto de esta investigación y, para lo cual el mecanismo propuesto no fue entrenado. Por otro lado, la confiabilidad del prototipo es

determinada mediante la evaluación de la precisión del mecanismo de estimación la cual fue evaluada y discutida ampliamente en el capítulo 4.

5.4. Implantación del prototipo

El despliegue del prototipo software se realizó sobre una estación con las siguientes características: procesador Core i3, Memoria RAM de 4G, Windows 10 y Disco duro de 500 GB, características encontradas en un computador personal promedio.

5.4.1. Vista de despliegue

A continuación, se describe la distribución física de los componentes que interactúan con en la aplicación. Se identifican dos nodos que acceden a una base de datos en común. El nodo Servicio Windows representa el programa en segundo plano usando la cámara como dispositivo de entrada para obtener las imágenes faciales, de manera que se obtiene la edad a partir de esas imágenes y se restringe el contenido de acuerdo a las reglas almacenadas en la base de datos. El nodo de control parental representa el programa ejecutable por el administrador que permite configurar las reglas y monitorear los eventos lanzados desde el servicio Windows. No hay una interacción directa entre los dos nodos, puesto que el control parental no puede ser ejecutado por un usuario sin privilegios de administración.

En la construcción de esta aplicación se tuvieron en cuenta las siguientes tecnologías:

- Visual Studio 2017: entorno de desarrollo
- OpenCV 2.4.1: librería de procesamiento de imágenes
- WinpkFilter 3.2: librería de captura de paquetes de red.
- Universal Windows Platform: plataforma común de aplicaciones disponible en todos los dispositivos que se ejecutan en Windows 10.
- C#: lenguaje de programación del Ejecutable de control parental.
- C++: lenguaje de programación de la librería de estimación de edad y control de red (Servicio instalable).
- RapidMiner Studio 7.5: entorno para minería de datos.

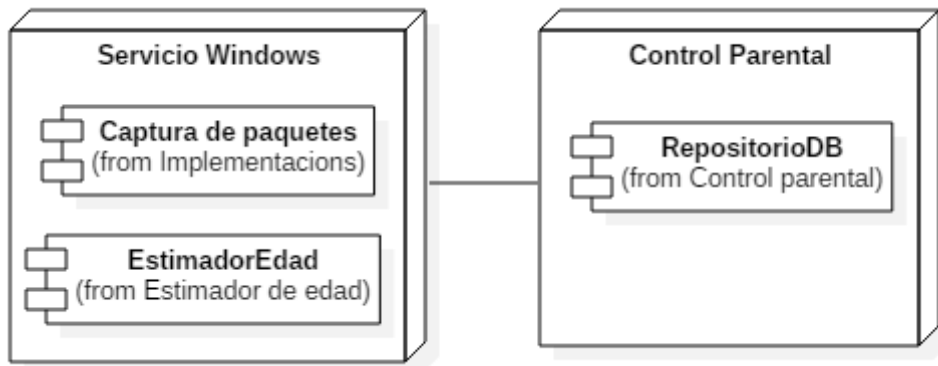


Figura 38. Nodos desplegados en un computador personal.

En la Figura 42 se muestran los nodos que integran el prototipo de control parental, el servicio de Windows es instalado en el sistema operativo y se ejecuta en segundo plano accediendo a la cámara del dispositivo, calculando la edad del usuario que se encuentra frente al computador. El nodo de control parental se instala como un programa ejecutable por el administrador del sistema operativo.

6. CONCLUSIONES Y TRABAJO FUTURO

6.1. Conclusiones

- Se presentó una amplia revisión de la literatura con respecto a los antecedentes y aspectos teóricos relacionados con los enfoques, tendencias y directrices, entre otros elementos para la estimación automática de la edad a partir de imágenes faciales, la cual permitió el desarrollo de este trabajo y podrá orientar futuras investigaciones en esta área.
- Se diseñó y construyó un nuevo mecanismo para la estimación automática de la edad a partir del procesamiento de imágenes faciales de los usuarios en tiempo real, que permite estimar un rango de edad al cual pertenecen.
- Se propuso una aplicación novedosa de apoyo a las funciones de control parental a nivel de prototipo funcional, a partir del mecanismo nuevo para la estimación de la edad por reconocimiento facial. Actualmente, no existen en el mercado este tipo de soluciones informáticas a partir de imágenes, que pueden ser usados para controlar el acceso a contenidos principalmente para menores de edad entre otras múltiples aplicaciones. El prototipo construido intercepta, clasifica y bloquea paquetes transmitidos mediante el protocolo de comunicación *http*.
- Fue determinado el mejor desempeño de la combinación de *Histogram Of Gradients* (HOG)+ *Local Binary Pattern* (LBP) e individualmente de HOG con respecto a LBP y *Compound Local Binary Pattern* (CLBP), como técnicas de descripción de imágenes faciales que permiten la estimación de la edad en grupos. Adicionalmente, en cuanto a clasificadores, *Support Vector Machine* (SVM) presentó un mejor desempeño con respecto a *k-Nearest Neighbors* (KNN), *Random Forest* (RF) y *Linear Discriminant Analysis* (LDA).
- Se consiguió mejorar en el mecanismo propuesto, la precisión en la estimación automática de la edad en grupos, evidenciándose esto a través de la aplicación bajo el mismo protocolo de evaluación y obteniéndose como resultados un *Accuracy of an Exact Match* (AEM) de 60.1% y *Accuracy of allowing an Error of One age category* (AEO) de 91.62% en el mecanismo propuesto por este proyecto con respecto a los valores de referencias más altos alcanzados de 56% y 88.8% respectivamente en los antecedentes revisados.
- El mecanismo de estimación de edad fue diseñado e implementado como un componente reutilizable para que pueda ser integrado en otros sistemas. Dicho mecanismo fue integrado en este caso con un sistema de control parental, sin embargo, el mecanismo puede ser adoptado para otros propósitos.
- Se diseñó y construyó un servicio de control parental el cual puede de ser ejecutado en segundo plano, acceder y filtrar los paquetes de red y, usar la cámara web del

ordenador para tomar sus decisiones de filtrado a partir del procesamiento de la imagen obtenida.

- La evaluación de confiabilidad del prototipo propuesta como uno de los objetivos de este trabajo fue determinada y realizada mediante la evaluación del mecanismo de estimación de edad, ampliamente discutido en el capítulo 4, dado que este mecanismo es el elemento central funcional del prototipo implementado.

6.2. Trabajo futuro

- Estudios relacionados con la inclusión de nuevas funcionalidades en el prototipo construido relacionadas con:
 - La detección de intentos de hackeo por medio de fotos frente a la cámara o uso de máscaras.
 - Ampliación de bases de datos para la categorización de las páginas web.
 - Módulo de administración web para permitir la centralización de la configuración para múltiples dispositivos y además de forma remota, ya integrados en otros sistemas de control parental consultados.
 - Análisis e inserción de los paquetes transmitidos mediante otros protocolos de comunicación.
- Evaluación del comportamiento del mecanismo de estimación de edad propuesto sobre otras bases de datos, como por ejemplo la base de datos MORPH o FG-Net.
- Investigar sobre otros descriptores como el descriptor global AAM, que podrían ser utilizados en el mecanismo automático de estimación de edad como posibles mejoras del nivel de estimación.
- Propiciar la construcción de una base de datos con las características ideales que según los antecedentes deberían cumplirse para la estimación de la edad a partir de imágenes faciales. Es decir, cubriendo un mayor rango de edad, desde recién nacidos hasta ancianos e incluyendo imágenes de un alto número de personas, mínimo 100, donde se cuente con al menos una imagen de cada persona por cada edad, idealmente almacenar una imagen de cada año desde el nacimiento hasta la vejez de cada persona.

7. REFERENCIAS BIBLIOGRÁFICAS

- [1] H. Han, C. Otto, A. K. Jain, and E. Lansing, "Age Estimation from Face Images: Human vs . Machine Performance," *Biometrics (ICB), 2013 Int. Conf.*, 2013.
- [2] L. A. Zebrowitz, *Reading Faces: Window to the Soul?*, no. 1. Westview Press, 1997.
- [3] Y. Fu, G. Guo, and T. S. Huang, "Age synthesis and estimation via faces: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 11, pp. 1955–1976, 2010.
- [4] G. Guo, Y. Fu, T. S. Huang, and C. R. Dyer, "Locally adjusted robust regression for human age estimation," *2008 IEEE Work. Appl. Comput. Vision, WACV*, 2008.
- [5] G. Guo, Y. Fu, C. R. Dyer, and T. S. Huang, "Image-based human age estimation by manifold learning and locally adjusted robust regression," *IEEE Trans. Image Process.*, vol. 17, no. 7, pp. 1178–1188, 2008.
- [6] U. Carlsson, "Violence and Pornography in the Media," *Int. Clear. Child.*, pp. 135–154, 2006.
- [7] C. A. Anderson, N. Ichori, B. J. Bushman, H. R. Rothstein, A. Shibuya, E. L. Swing, A. Sakamoto, and M. Saleem, "Violent video game exposure and aggression: A literature review," *Minerva Psichiatr.*, vol. 45, no. 1, pp. 1–18, 2004.
- [8] B. Nelson, G. H. Smith, and J. Ensign, "The impact of media violence on children on committee on commerce," pp. 1–120, 2012.
- [9] "Children and the media," *Paediatr. Child Health*, vol. 4, no. 5, pp. 350–354, 1999.
- [10] "How does Netflix decide the maturity rating on movies and TV shows?" [Online]. Available: <http://goo.gl/6VAwvB>. [Accessed: 28-Apr-2016].
- [11] C. De Colombia, "Ley 1098 de 2006 1/118," vol. 2006, no. 46. pp. 1–118, 2007.
- [12] UNICEF, "Convención Sobre Los Derechos Del Niño," 2015.
- [13] A. Lanitis, C. Draganova, and C. Christodoulou, "Comparing different classifiers for automatic age estimation," *IEEE Trans. Syst. Man, Cybern. Part B*, vol. 34, no. 1, pp. 621–628, Feb. 2004.
- [14] "Releyeable." [Online]. Available: <http://goo.gl/7UGST9>. [Accessed: 28-Apr-2016].
- [15] Cognitech, "FaceVACS-VideoScan, Technical Specification," 2015.
- [16] "Net Nanny," 2016. [Online]. Available: <http://goo.gl/Pzf1mM>. [Accessed: 23-May-2016].
- [17] "Play store. Qustodio Control Parental." [Online]. Available: <http://goo.gl/dQVPGo>. [Accessed: 28-Apr-2016].
- [18] "Norton Family." [Online]. Available: <http://goo.gl/RnZBPp>. [Accessed: 19-May-2016].
- [19] "Qustodio." [Online]. Available: <http://goo.gl/FsBTW4>. [Accessed: 19-May-2016].
- [20] "McAfee Parental Controls." [Online]. Available: <http://goo.gl/t4xvMZ>. [Accessed: 04-May-

2016].

- [21] S. W. Ambler, "The Agile Unified Process," 2016. [Online]. Available: <http://goo.gl/ayowrg>. [Accessed: 23-May-2016].
- [22] A. de la Escalera Hueso, *Visión por computador: fundamentos y métodos*. Pearson Educación, 2001.
- [23] T. Huang, Z. Xiong, and Z. Zhang, "Face Recognition Applications," in *Handbook of Face Recognition*, S. Z. Li and A. K. Jain, Eds. Springer-Verlag London, 2011, pp. 617–638.
- [24] P. Carcagnì, M. Del Coco, D. Cazzato, M. Leo, and C. Distante, "A study on different experimental configurations for age, race, and gender estimation problems," *EURASIP J. Image Video Process.*, vol. 2015, no. 1, p. 37, 2015.
- [25] C. Information and Privacy Commissioner of Ontario, "Anonymous Video Analytics (AVA) technology and privacy," 2011.
- [26] C. Chen, A. Dantcheva, and A. Ross, "Impact of facial cosmetics on automatic gender and age estimation algorithms," *Comput. Vis. Theory Appl. (VISAPP), 2014 Int. Conf.*, vol. 2, pp. 182–190, 2014.
- [27] G. Guo, "Human Age Estimation and Sex Classification," in *Video Analytics for Business Intelligence*, C. Shan, F. Porikli, T. Xiang, and S. Gong, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 101–131.
- [28] G. Bradski and A. Kaehler, *Learning OpenCV: Computer Vision in C++ with the OpenCV Library*, 2nd ed. O'Reilly Media, Inc., 2013.
- [29] G. Panis, A. Lanitis, N. Tsapatsoulis, and T. F. Cootes, "Overview of research on facial ageing using the FG-NET ageing database," *IET Biometrics*, vol. 5, no. 2, pp. 37–46, 2016.
- [30] O. García-Olalla and E. Alegre, "Descripción de textura en imágenes utilizando Local Binary Pattern (LBP)," in *Conceptos y métodos en visión por computador*, 2016, pp. 115–130.
- [31] P. Sharma and M. Kaur, "Classification in Pattern Recognition: A Review," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 3, no. 4, pp. 2277–128, 2013.
- [32] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.
- [33] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A Training Algorithm for Optimal Margin Classifiers," *Proc. fifth Annu. Work. Comput. Learn. theory*, pp. 144–152, 1992.
- [34] Carmona E., "Tutorial sobre Máquinas de Vectores Soporte (SVM)," 2014.
- [35] G. Betancour, "LAS MÁQUINAS DE SOPORTE VECTORIAL (SVMs)," *Sci. Tech.*, no. 27, pp. 67–72, 2005.
- [36] G. Guo, G. Mu, Y. Fu, C. Dyer, and T. Huang, "A study on automatic age estimation using a large database," *Comput. Vision, 2009 IEEE 12th Int. Conf.*, vol. 12, no. C, pp. 1986–1991, 2009.
- [37] C. M. Bishop, "Neural Networks for Pattern Recognition," *Oxford Univ. Press*, 1996.
- [38] V. N. Vapnik, "Statistical Learning Theory," *John Wiley*, 1998.

- [39] C. Fernández, I. Huerta, and A. Prati, "A comparative evaluation of regression learning algorithms for facial age estimation," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8912, pp. 133–144, 2015.
- [40] K. Chang, C. Chen, and Y. Hung, "A ranking approach for human age estimation based on face," *Proc. Int'l Conf. Pattern Recognit.*, 2010.
- [41] Y. H. Kwon and V. Lobo, "Age Classification from Facial Images," vol. 74, no. 1, pp. 1–21, 1999.
- [42] K. H. Liu, S. Yan, and C. C. J. Kuo, "Age group classification via structured fusion of uncertainty-driven shape features and selected surface features," *2014 IEEE Winter Conf. Appl. Comput. Vision, WACV 2014*, pp. 445–452, 2014.
- [43] S. Izadpanahi and Ö. Toygar, "Geometric feature based age classification using facial images," *IET Conf. Image Process. (IPR 2012)*, pp. 1–5, 2012.
- [44] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active Appearance Models," *Proc. Eur. Conf. Comput. Vis.*, vol. 2, pp. 484–498, 1998.
- [45] L. Liu, J. Liu, and J. Cheng, "Age-group classification of facial images," *Proc. - 2012 11th Int. Conf. Mach. Learn. Appl. ICMLA 2012*, vol. 1, pp. 693–696, 2012.
- [46] S. E. Choi, Y. J. Lee, S. J. Lee, K. R. Park, and J. Kim, "Age estimation using a hierarchical classifier based on global and local facial features," *Pattern Recognit.*, vol. 44, no. 6, pp. 1262–1281, 2011.
- [47] W. Li, Y. Wang, and Z. Zhang, "A hierarchical framework for image-based human age estimation by weighted and OHRanked Sparse Representation-based classification," *Proc. - 2012 5th IAPR Int. Conf. Biometrics, ICB 2012*, pp. 19–25, 2012.
- [48] P. X. Gao, "Facial age estimation using Clustered Multi-task Support Vector Regression Machine," *Proc. - Int. Conf. Pattern Recognit.*, no. 1cpr, pp. 541–544, 2012.
- [49] X. Geng, Z. Zhou, Y. Zhang, G. Li, and H. Dai, "Learning from Facial Aging Patterns for Automatic Age Estimation," *Proc. 14th Annu. ACM Int. Conf. Multimed. - Multimed. '06*, 2006.
- [50] M. Bereta, P. Karczmarek, W. Pedrycz, and M. Reformat, "Local descriptors in application to the aging problem in face recognition," *Pattern Recognit.*, vol. 46, no. 10, pp. 2634–2646, 2013.
- [51] C. Chen, W. Yang, Y. Wang, S. Shan, and K. Ricanek, "Learning Gabor Features for Facial Age Estimation," in *Biometric Recognition: 6th Chinese Conference, CCBR 2011, Beijing, China, December 3-4, 2011. Proceedings*, Z. Sun, J. Lai, X. Chen, and T. Tan, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 204–213.
- [52] C. Chen, W. Yang, Y. Wang, K. Ricanek, and K. Luu, "Facial feature fusion and model selection for age estimation," *Face Gesture 2011*, pp. 200–205, 2011.
- [53] X. Sun, P. Wu, and H. Liu, "Facial Age Estimation Using Bio-Inspired Features and Cost-Sensitive Ordinal Hyperplane Rank," pp. 81–85, 2014.
- [54] D. Yi, Z. Lei, and S. Z. Li, "Age estimation by multi-scale convolutional network," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9005, pp. 144–158, 2015.

- [55] J. Qiu, Y. Dai, Y. Zhang, and J. M. Alvarez, "Hierarchical Aggregation Based Deep Aging Feature for Age Prediction," *2015 Int. Conf. Digit. Image Comput. Tech. Appl. DICTA 2015*, 2016.
- [56] G. Levi and T. Hassner, "Age and Gender Classification using Convolutional Neural Networks," pp. 34–42, 2015.
- [57] R. Ranjan, S. Zhou, and J. C. Chen, "Unconstrained Age Estimation with Deep Convolutional Neural Networks," pp. 109–117, 2015.
- [58] K. Y. Chang, C. S. Chen, and Y. P. Hung, "Ordinal hyperplanes ranker with cost sensitivities for age estimation," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 585–592, 2011.
- [59] K. Y. Chang and C. S. Chen, "A learning framework for age rank estimation based on face images with scattering transform," *IEEE Trans. Image Process.*, vol. 24, no. 3, pp. 785–798, 2015.
- [60] R. Weng, J. Lu, G. Yang, and Y. P. Tan, "Multi-feature ordinal ranking for facial age estimation," *2013 10th IEEE Int. Conf. Work. Autom. Face Gesture Recognition, FG 2013*, 2013.
- [61] Y. Zhang, "Age Estimation Using Bayesian Process," in *New Frontiers in Applied Data Mining: PAKDD 2011 International Workshops, Shenzhen, China, May 24-27, 2011, Revised Selected Papers*, L. Cao, J. Z. Huang, J. Bailey, Y. S. Koh, and J. Luo, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 447–458.
- [62] Y. Su, M. Wang, and H. Bao, "Age-variation face recognition based on bayes inference," *2015 12th Int. Conf. Fuzzy Syst. Knowl. Discov. FSKD 2015*, pp. 1108–1112, 2016.
- [63] D. Gao, L. Pan, R. Liu, R. Chen, and M. Xie, "CORRELATED WARPED GAUSSIAN PROCESSES FOR GENDER-SPECIFIC AGE ESTIMATION University of Electronic Science and Technology of China School of Software Technology , Dalian University of Technology State Key Lab . for Novel Software Technology , Nanjing Univer," vol. 1, pp. 133–137, 2015.
- [64] X. Geng, C. Yin, and Z. H. Zhou, "Facial age estimation by learning from label distributions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 10, pp. 2401–2412, 2013.
- [65] C. Yin and X. Geng, "Facial Age Estimation by Conditional Probability Neural Network," pp. 243–250, 2012.
- [66] K. H. Liu, S. Yan, and C. C. J. Kuo, "Age Estimation via Grouping and Decision Fusion," *IEEE Trans. Inf. Forensics Secur.*, vol. 10, no. 11, pp. 2408–2423, 2015.
- [67] J. Si, J. Feng, Q. Bu, X. Sun, X. He, and S. Qiu, "Age Estimation Based on Canonical Correlation Analysis and Extreme Learning Machine," in *Biometric Recognition: 10th Chinese Conference, CCBR 2015, Tianjin, China, November 13-15, 2015, Proceedings*, J. Yang, J. Yang, Z. Sun, S. Shan, W. Zheng, and J. Feng, Eds. Cham: Springer International Publishing, 2015, pp. 677–685.
- [68] J. K. Pontes, A. S. Britto, C. Fookes, and A. L. Koerich, "A flexible hierarchical approach for facial age estimation based on multiple features," *Pattern Recognit.*, pp. 1–18, 2016.
- [69] N. Borzue and K. Faez, "Facial Age Estimation Using Artificial Neural Network and Robust Feature Extraction," no. November, pp. 25–27, 2015.

- [70] R. Tokola, D. Bolme, C. Boehnen, D. Barstow, and K. Ricanek, "Discriminating projections for estimating face age in wild images," *IJCB 2014 - 2014 IEEE/IAPR Int. Jt. Conf. Biometrics*, 2014.
- [71] J. Ylioinas, A. Hadid, X. Hong, and M. Pietikäinen, "Age Estimation Using Local Binary Pattern Kernel Density Estimate," in *Image Analysis and Processing -- ICIAP 2013: 17th International Conference, Naples, Italy, September 9-13, 2013. Proceedings, Part I*, A. Petrosino, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 141–150.
- [72] T. Wu, P. Turaga, and R. Chellappa, "Age estimation and face verification across aging using landmarks," *IEEE Trans. Inf. Forensics Secur.*, vol. 7, no. 6, pp. 1780–1788, 2012.
- [73] L. Prevost, P. Phothisane, and E. Bigorgne, "Real-Time Facial Analysis in Still Images and Videos for Gender and Age Estimation," in *Pattern Recognition Applications and Methods: Third International Conference, ICPRAM 2014, Angers, France, March 6-8, 2014, Revised Selected Papers*, A. Fred, M. De Marsico, and A. Tabbone, Eds. Cham: Springer International Publishing, 2015, pp. 263–278.
- [74] C. Hu, L. Gong, T. Wang, and Q. Feng, "Effective human age estimation using a two-stage approach based on Lie Algebrized Gaussians feature," *Multimed. Tools Appl.*, vol. 74, no. 11, pp. 4139–4159, 2015.
- [75] J. Lu, V. E. Liong, and J. Zhou, "Cost-Sensitive Local Binary Feature Learning for Facial Age Estimation," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5356–5368, 2015.
- [76] J. M. Guo, Y. M. Liou, and H. S. Nguyen, "Human face age estimation with adaptive hybrid features," *Proc. 2011 Int. Conf. Syst. Sci. Eng.*, no. June, pp. 55–58, 2011.
- [77] H. Han, C. Otto, X. Liu, and A. K. Jain, "Demographic estimation from face images: Human vs. machine performance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 6, pp. 1148–1161, 2015.
- [78] M. Y. El Dib and H. M. Onsi, "Human age estimation framework using different facial parts," *Egypt. Informatics J.*, vol. 12, no. 1, pp. 53–59, 2011.
- [79] M. Y. El Dib and M. El-saban, "HUMAN AGE ESTIMATION USING ENHANCED BIO-INSPIRED FEATURES (EBIF) Faculty of Computers and Information , Cairo University , Cairo , Egypt," pp. 1–4, 2010.
- [80] O. F. W. Onifade and J. D. Akinyemi, "GWAgeER – A GroupWise Age Ranking Framework for Human Age Estimation," *Int. J. Image Graph. Signal Process.*, vol. 7, no. 5, pp. 1–12, 2015.
- [81] T. Gehrig, M. Steiner, and H. K. Ekenel, "Draft: Evaluation Guidelines for Gender Classification and Age Estimation," pp. 1–12, 2011.
- [82] X. Geng, Q. Wang, and Y. Xia, "Facial age estimation by adaptive label distribution learning," *Proc. - Int. Conf. Pattern Recognit.*, pp. 4465–4470, 2014.
- [83] I. Huerta, C. Fernández, and A. Prati, "Facial Age Estimation Through the Fusion of Texture and Local Appearance Descriptors," in *Computer Vision - ECCV 2014 Workshops: Zurich, Switzerland, September 6-7 and 12, 2014, Proceedings, Part II*, L. Agapito, M. M. Bronstein, and C. Rother, Eds. Cham: Springer International Publishing, 2015, pp. 667–681.
- [84] C. Shan, "Learning local features for age estimation on real-life faces," *Proc. 1st ACM Int. Work. Multimodal pervasive video Anal. - MPVA '10*, p. 23, 2010.

- [85] A. Gallagher and T. Chen, "Understanding Images of Groups of People," in *Proc. CVPR*, 2009.
- [86] A. C. Gallagher and T. Chen, "Understanding images of groups of people," *2009 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work. CVPR Work. 2009*, pp. 256–263, 2009.
- [87] C. Li, Q. Liu, J. Liu, and H. Lu, "Learning ordinal discriminative features for age estimation," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 2570–2577.
- [88] J. Ylioinas, A. Hadid, and M. Pietikainen, "Age Classification in Unconstrained Conditions Using LBP Variants," *2012 21st Int. Conf. Pattern Recognit. (Icpr 2012)*, no. Icpr, pp. 1257–1260, 2012.
- [89] S. E. Bekhouche and A. Ouafi, "AUTOMATIC AGE ESTIMATION AND GENDER CLASSIFICATION IN THE WILD Laboratory of LAGE , University of Ouargla , Algeria LAMIH Laboratory , UMR CNRS 8201 UVHC , University of Valenciennes , France Center for Machine Vision Research , University of Oulu , Finla," 2015.
- [90] S. E. Bekhouche, A. Ouafi, A. Benlamoudi, A. Taleb-Ahmed, and A. Hadid, "Facial age estimation and gender classification using multi level local phase quantization," *3rd Int. Conf. Control. Eng. Inf. Technol. CEIT 2015*, pp. 3–6, 2015.
- [91] "Microsoft cognitives services. Computer Visio API." [Online]. Available: <http://goo.gl/M4TXRy>. [Accessed: 28-Apr-2016].
- [92] D. Fechtig, "Informe de producto ReleEYEble," vol. 49, no. 0, pp. 1–5, 2008.
- [93] AMSCREEN, "Indoor and Outdoor Screens," *Amscreen.eu*, 2016. [Online]. Available: <http://goo.gl/emWgWT>. [Accessed: 18-May-2016].
- [94] NEC Display Solutions United Kingdom, "NEC FieldAnalyst," 2016. [Online]. Available: <http://goo.gl/GfKhY7>. [Accessed: 18-May-2016].
- [95] "Mobile spy," 2016. [Online]. Available: <http://goo.gl/070UPe>. [Accessed: 23-May-2016].
- [96] "Phone Sheriff," 2016. [Online]. Available: <http://goo.gl/tEVWIZ>. [Accessed: 23-May-2016].
- [97] "Mob stealth," 2016. [Online]. Available: <http://goo.gl/AKzVck>. [Accessed: 23-May-2016].
- [98] "Auto Forward," 2016. [Online]. Available: <http://goo.gl/U1oU5H>. [Accessed: 23-May-2016].
- [99] "Pumpic," 2016. [Online]. Available: <http://goo.gl/fl5o0K>. [Accessed: 23-May-2016].
- [100] "Easy Spy," 2016. [Online]. Available: <http://goo.gl/2UN0Uh>. [Accessed: 23-May-2016].
- [101] "Witigo," 2016. [Online]. Available: <http://goo.gl/8Tu1OC>. [Accessed: 23-May-2016].
- [102] "Web Watcher," 2016. [Online]. Available: <http://goo.gl/N1e23g>. [Accessed: 23-May-2016].
- [103] "Internet Safety," 2016. [Online]. Available: <http://goo.gl/yEjDDs>. [Accessed: 23-May-2016].
- [104] "Intego," 2016. [Online]. Available: <http://goo.gl/chrqj6>. [Accessed: 23-May-2016].
- [105] "Spy Agent," 2016. [Online]. Available: <http://goo.gl/JcSUJ4>. [Accessed: 23-May-2016].
- [106] "Cyber Sitter," 2016. [Online]. Available: <http://goo.gl/7iKR1b>. [Accessed: 23-May-2015].

- [107] "Nch Software," 2016. [Online]. Available: <http://goo.gl/4ix9Eb>. [Accessed: 23-May-2016].
- [108] "Elite Keyloggers," 2016. [Online]. Available: <http://goo.gl/BbnsFU>. [Accessed: 23-May-2016].
- [109] "Facelt." [Online]. Available: <http://goo.gl/rhM15y>. [Accessed: 19-May-2016].
- [110] "Omron." [Online]. Available: <http://goo.gl/U6igf8>. [Accessed: 19-May-2016].
- [111] "Face Key," 2014. [Online]. Available: <http://goo.gl/HU8wxt>. [Accessed: 19-May-2016].
- [112] Cognitec, "FaceVACS-Entry," 2016. [Online]. Available: <http://goo.gl/1fWkX7>. [Accessed: 19-May-2016].
- [113] D. E. King, "Max-Margin Object Detection," 2015.
- [114] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *Proc. 2001 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognition. CVPR 2001*, vol. 1, p. I-511-I-518, 2001.
- [115] T. Baltrušaitis, P. Robinson, and L. P. Morency, "Constrained local neural fields for robust facial landmark detection in the wild," *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 354–361, 2013.
- [116] D. Cristinacce and T. F. Cootes, "Feature Detection and Tracking with Constrained Local Models," *Proceedings Br. Mach. Vis. Conf. 2006*, p. 95.1-95.10, 2006.
- [117] T. Baltrusaitis, P. Robinson, and L. P. Morency, "OpenFace: An open source facial behavior analysis toolkit," *2016 IEEE Winter Conf. Appl. Comput. Vision, WACV 2016*, 2016.
- [118] T. Ojala, M. Pietikainen, and D. Harwood, "Performance evaluation of texture measures with classification based on Kullback discrimination of distributions," *Pattern Recognition, 1994. Vol. 1 - Conf. A Comput. Vis. amp; Image Process. Proc. 12th IAPR Int. Conf.*, vol. 1, pp. 582–585 vol.1, 1994.
- [119] Y. Yang and S. Newsam, "Comparing sift descriptors and gabor texture features for classification of remote sensed imagery," in *Proceedings - International Conference on Image Processing, ICIP, 2008*, pp. 1852–1855.
- [120] F. Ahmed, E. Hossain, A. S. M. H. Bari, and A. Shihavuddin, "Compound local binary pattern (CLBP) for robust facial expression recognition," *12th IEEE Int. Symp. Comput. Intell. Informatics, CINTI 2011 - Proc.*, no. December 2016, pp. 391–395, 2011.
- [121] T. Jabid, M. H. Kabir, and O. Chae, "Robust facial expression recognition based on local directional pattern," *ETRI J.*, vol. 32, no. 5, pp. 784–794, 2010.
- [122] M. P. and T. M. Timo Ojala, "Multiresolution Gray Scale and Rotation Invariant Texture Classification with Local Binary Patterns Timo," *Proc. - Int. Conf. Image Process. ICIP*, pp. 1852–1855, 2008.
- [123] P. Belhumeur, J. Hespanha, and D. Kriegman, "Eigenfaces Vs. Fisherfaces: Recognition Using Class Specific Linear Projection," vol. 19, no. 7, pp. 711–720, 1997.
- [124] R. Azarmehr, R. Lagani, W. Lee, C. Xu, and D. Laroche, "Real-time Embedded Age and Gender Classification in Unconstrained Video," pp. 56–64, 2015.
- [125] C. Chang and C. Lin, "LIBSVM: A Library for Support Vector Machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, pp. 1–39, 2013.

- [126] and C.-J. L. Chih-Wei Hsu, Chih-Chung Chang, "A Practical Guide to Support Vector Classification," *BJU Int.*, vol. 101, no. 1, pp. 1396–400, 2008.
- [127] A. Torrisi, G. M. Farinella, G. Puglisi, and S. Battiato, "SELECTING DISCRIMINATIVE CLBP PATTERNS FOR AGE ESTIMATION IPLAB", Department of Mathematics and Computer Science, University of Catania Department of Mathematics and Computer Science, University of Cagliari."
- [128] G. Panis and A. Lanitis, "An Overview of Research Activities in Facial Age Estimation Using the FG-NET Aging Database," in *Computer Vision - ECCV 2014 Workshops: Zurich, Switzerland, September 6-7 and 12, 2014, Proceedings, Part II*, L. Agapito, M. M. Bronstein, and C. Rother, Eds. Cham: Springer International Publishing, 2015, pp. 737–750.
- [129] K. Ricanek and T. Tesafaye, "MORPH: A longitudinal image database of normal adult age-progression," *FGR 2006 Proc. 7th Int. Conf. Autom. Face Gesture Recognit.*, vol. 2006, no. September, pp. 341–345, 2006.
- [130] R. M. Borwankar, G. S. Pednekar, S. A. Deshpande, P. S. Sawant, and S. S. Chavan, "Comparative Evaluation of Age Classification from Facial Images," pp. 2244–2249, 2015.
- [131] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
- [132] S. Balakrishnama and a Ganapathiraju, "Linear Discriminant Analysis - a Brief Tutorial," *Compute*, vol. 11, pp. 1–9, 1998.
- [133] Panda Security, "Control Parental," 2016. [Online]. Available: <http://goo.gl/0aT0Hj>. [Accessed: 18-May-2016].
- [134] T. J. Shimeall and J. M. Spring, "Chapter 5 - Frustration Strategies: Technical Controls," in *Introduction to Information Security*, T. J. Shimeall and J. M. Spring, Eds. Boston: Syngress, 2014, pp. 83–106.
- [135] "SquidGuard." [Online]. Available: <https://goo.gl/FSDEPy>. [Accessed: 03-Dec-2017].
- [136] N. Kernel, "Windows Packet Filter," 2017. [Online]. Available: <https://goo.gl/MOkz4p>. [Accessed: 15-Feb-2017].
- [137] "Rapidminer Studio." RapidMiner Inc., 2017.
- [138] X. Geng, Z. Zhou, S. Member, K. Smith-miles, and S. Member, "Automatic Age Estimation Based on Facial Aging Patterns," *Pattern Anal. Mach. Intell. IEEE Trans.*, vol. 29, no. 12, pp. 2234–2240, 2007.
- [139] Roy Thomas Fielding, "Representational State Transfer (REST)," 2000. [Online]. Available: <http://goo.gl/SeMG4i>. [Accessed: 19-May-2016].
- [140] T. Berners-Lee, R. Fielding, and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax," 2005. [Online]. Available: <http://goo.gl/ALOZZV>. [Accessed: 19-May-2016].
- [141] W. Hyun, M. Huh, S. Kim, and S. Kang, "Study on design and implementation of audience measurement functionalities for digital signage service using Kinect camera," 2014.