



**UNIVERSIDAD  
DE ANTIOQUIA**

**Reconocimiento de sujetos y caracterización de sus  
emociones a partir de sus expresiones faciales usando  
algoritmos de reconocimiento facial**

Autor

**Mauricio Saldarriaga Quintero**

Universidad de Antioquia

Facultad de ingeniería, Programa de bioingeniería

Medellín, Colombia

2020



Reconocimiento de sujetos y caracterización de sus emociones a partir de sus expresiones  
faciales usando algoritmos de reconocimiento facial

**Mauricio Saldarriaga Quintero**

Práctica Académica en modalidad semestre de industria como requisito parcial para optar al

título de:

**Bioingeniero**

Asesora:

MSc. Jenny Kateryne Aristizábal Nieto

Universidad de Antioquia  
Facultad de Ingeniería  
Bioingeniería  
Medellín, Colombia  
2020

## **Resumen**

La empresa Fourier Tech se dedica a crear soluciones para la industria basados en modelos matemáticos complejos, una de sus áreas es la inteligencia artificial que se ocupa de crear herramientas con capacidades de aprendizaje similares a las del ser humano. Se requería diseñar un sistema de reconocimiento facial que pudiese ser usado en gran variedad de aplicaciones, en este caso se dio un enfoque a las áreas de la salud para que pudiese ser usado en situaciones de emergencia identificando sujetos. Se desarrolló un sistema funcional capaz de detectar un rostro en una imagen y extraer las características para realizar estimaciones sobre la edad, el género y la expresión facial. El sistema estaba asociado a una base de datos creada con varios sujetos de prueba, donde se almacenaban sus datos biométricos y sus datos personales con el fin de que estos fuesen retornados después de ser reconocidos. Se realizaron pruebas para evaluar la precisión del sistema a la hora de realizar estimaciones y finalmente se hicieron pruebas de desempeño con las que se evaluó la calidad del sistema desde el punto de vista funcional, de diseño y de uso.

## Tabla de contenido

1. Introducción.....	5
2. Objetivos.....	6
2.1 Objetivo general.....	6
2.2 Objetivos específicos.....	6
3. Marco teórico .....	7
3.1 Inteligencia artificial .....	7
3.2 Reconocimiento facial .....	7
3.3 Antecedentes y estado del arte .....	8
3.4 Redes neuronales.....	9
3.4.1 Redes neuronales convolucionales .....	10
3.5 Bases de datos .....	11
3.6 JavaScript y TensorFlow .....	11
3.6.1 Faceapi.....	11
4. Metodología .....	14
4.1 Revisión bibliográfica .....	14
4.2 Búsqueda de librerías para detección de rostros en JavaScript y Tensorflow .....	14
4.3 Desarrollo de la interfaz gráfica de usuario.....	14
4.4 Diseño de la base de datos .....	15
4.5 Validación de modelos para identificación expresiones faciales, edad y género.....	15
4.6 Diseño del sistema de reconocimiento en tiempo real.....	16
4.7 Registros de datos biométricos de rostros con su respectiva información.....	17
4.8 Diseño de pruebas de desempeño.....	17
4.9 Diseño de pruebas para evaluar identificación de expresiones faciales, edad y género. ....	19
5. Resultados .....	20
5.1 Detección.....	20
5.2 Puntos de referencia faciales, edad y género.....	20
5.3 Interfaz gráfica .....	20
5.4 Base de datos.....	22
5.5 Expresiones faciales .....	23
5.6 Reconocimiento en tiempo real.....	24

5.7 Pruebas de reconocimiento editando los registros .....	24
5.8 Pruebas de desempeño.....	25
5.9 Validación de parámetros.....	26
6. Análisis de resultados.....	28
6.1 Interfaz gráfica y reconocimiento .....	28
6.2 Base de datos .....	28
6.3 Pruebas de reconocimiento al realizar modificaciones.....	29
6.4 Pruebas de desempeño.....	30
6.5 Expresiones faciales .....	31
6.6 Edad y género .....	32
7. Conclusiones.....	33
Bibliografía .....	34

## 1. Introducción

Existen diversas situaciones en el área de la salud que requieren la identificación rápida de información básica de sujetos que se encuentran en una situación de urgencia, datos como el tipo de sangre, edad, alergias y algunos antecedentes pueden ser importantes a la hora de atender una emergencia [1]. Sujetos que llegan desorientados a centros de salud, personas con enfermedades neurodegenerativas que son incapaces de identificarse, y en general todos los casos en los que haya problemas de identificación hacen surgir una pregunta: ¿Cómo pueden las nuevas tecnologías brindar una herramienta que permita manejar situaciones que requieran reconocer la identidad y algunos datos de sujetos en condiciones de emergencia?

Actualmente los sistemas de autenticación juegan un papel fundamental en situaciones cotidianas, son usados en universidades, fabricas, transacciones en línea, hospitales, entre muchos otros. Los métodos tradicionales para abordar el problema se basan en documentos de identidad, códigos, usuarios y contraseñas. sin embargo, no son muy seguros y son poco prácticos, es por eso que la biometría es una alternativa cada vez más usada, realiza la tarea de identificación a través de las características fisiológicas humanas de forma automática. Algunas de las métricas más usadas son la huella dactilar, iris, geometría de la mano, reconocimiento de voz y de patrones faciales [2]. Los sistemas que utilizan biometría para la identificación tienen la desventaja de requerir un sensor para la obtención de los datos, en los sistemas que usan reconocimiento facial se tiene la ventaja de fácil adquisición por medio de cámaras que no interactúan directamente con los usuarios y que pueden estar en cualquier dispositivo móvil [3].

Existen diversas herramientas de reconocimiento facial basadas en algoritmos de inteligencia artificial que pueden reconocer patrones por medio de redes neuronales, las cuales son entrenadas con miles de imágenes para obtener una alta precisión [4]. Actualmente estas herramientas están disponibles y se integran en otros códigos para crear aplicaciones y no tener que resolver el problema desde el principio.

El aplicativo que se desarrolla en este proyecto contiene los algoritmos para la detección de rostro, mediante el cual se ubica al sujeto frente a una cámara y se reconoce su identidad en caso de ya estar registrado en la base de datos, o en su defecto se hace el debido ingreso y registro de este. Se utilizan algunas librerías de TensorFlow que es una plataforma abierta para desarrolladores con diversos modelos entrenados para obtener patrones de reconocimiento facial y extracción de características usadas para la estimación de la edad, el género y las expresiones faciales [5]. Finalmente se utiliza el lenguaje de programación JavaScript a fin de crear una interfaz de usuario para el manejo del sistema diseñado.

## **2. Objetivos**

### **2.1 Objetivo general**

Desarrollar un sistema de reconocimiento facial e identificación de emociones, asociado a una base de datos que contenga información personal de sujetos registrados.

### **2.2 Objetivos específicos**

- 1.** Desarrollar un algoritmo de reconocimiento facial que permita identificar personas y relacionar sus datos personales.
- 2.** Diseñar una base de datos registrando sujetos de prueba y sus respectivos datos.
- 3.** Entrenar una red neuronal que permita diferenciar emociones a partir de las expresiones faciales de los sujetos.
- 4.** Diseñar una interfaz gráfica que permita usar una cámara para el ingreso de una imagen y retorne los datos asociados a un sujeto después de identificarlo en tiempo real.
- 5.** Validar el funcionamiento del sistema de reconocimiento por medio de pruebas y registros.

### 3. Marco teórico

#### 3.1 Inteligencia artificial

Existen muchas definiciones de inteligencia artificial y suelen variar con cada autor, de forma general se puede decir que se trata de una disciplina científica que desarrolla programas con habilidades propias de los seres humanos tales como el aprendizaje y el razonamiento lógico [6]. Es importante diferenciar entre un programa común y uno con inteligencias artificial, para esto hay un método que es el test de Turing que consiste en poner a interactuar un humano con el ordenador por medio de una conversación, si el humano es incapaz de diferenciar si esta interactuando con un humano o con un ordenador, entonces el programa aprueba el test y se dice que posee inteligencia artificial [7].

Andreas Kaplan y Michael Haenlein definen la inteligencia artificial como un sistema que interpreta datos y los usa para lograr objetivos a través de la adaptación [8]. En este proyecto, la inteligencia artificial se usa para el reconocimiento de rostros humanos de modo automático por parte de un computador, que es la forma de interpretación de datos por medio de la cual se quiere llegar a la información de un sujeto.

La **visión artificial** es una disciplina derivada de la inteligencia artificial y consiste en el desarrollo de sistemas incluyendo hardware y software que a partir del procesamiento de imágenes pueden brindar potentes herramientas en la industria, academia, investigación y en cualquier campo que se requiera, los procesos de visión artificial consisten en diferenciar objetos y formas en una imagen y son posibles gracias a tecnologías basadas en la captura de la imagen (cámaras de vídeo, cámaras web) sumadas a la capacidad de procesamiento de los ordenadores [9].

#### 3.2 Reconocimiento facial

El reconocimiento facial hace parte de la visión artificial, el rostro humano posee gran cantidad de información sobre un sujeto y nos permite diferenciar a simple vista entre sujetos a partir de sus rasgos, los más significativos son los ojos, las orejas, las cejas, la nariz y la boca; a partir de la detección de estos en una imagen se puede hacer una caracterización biométrica y utilizar estos datos para que un ordenador logre detectar y reconocer un rostro en una imagen [10].

Por lo general los sistemas de reconocimiento facial se componen de una estructura similar, la primera etapa es la **detección** que consiste en la implementación de un algoritmo que diferencie un rostro en una imagen, para esto la convierte a blanco y negro para obtener un histograma de gradientes orientados que son vectores que indican las diferencias de brillo entre pixeles para detectar los bordes en la imagen, después esta la etapa de **alineación** en la que el sistema determina el tamaño y la posición de la cabeza, seguido a esto se encuentra la **normalización** en la que se estira y deforma la imagen con el fin de orientar el rostro para que mire hacia el frente y se pone a escala para un adecuado procesamiento, la siguiente etapa

es la de **representación** donde se unifica el código en función de la extracción biométrica que se realice del rostro y finalmente el proceso de **comparación** con otros rostros para establecer la similitud [11].

### 3.3 Antecedentes y estado del arte

El inicio de los sistemas de reconocimiento facial se puede ubicar en la década de los 60 cuando W. W. Bledsoe junto con su equipo de investigación desarrollan un sistema que no era automático ya que requería de un operador que ubicara los rasgos característicos en los rostros [12], en los años 70 algunos investigadores intentaron automatizar el proceso, Harmon y Kanade lograron implementar el sistema capaz de reconocer caras humanas por medio de una computadora identificando perfiles de rostros dibujados por un artista [13], Posteriormente Kanade en su tesis doctoral desarrollo un sistema de reconocimiento de alta precisión logrando detectar 608 rostros en una base de datos con 670 imágenes [14].

En 1988 L. Sirobich y M. Kirby hicieron uso de la técnica de análisis de componentes principales (PCA), que es una técnica de álgebra lineal y la usaron como método de automatización para el reconocimiento de rostros con lo que pudieron demostrar que bastaba con menos de 100 componentes para cifrar adecuadamente un rostro [15].

Los sistemas de reconocimiento facial fueron mejorando con los años y cada vez se perfeccionaban las técnicas para buscar optimizar el coste computacional del procesamiento de las imágenes y mejorar el rendimiento del software, un ejemplo de esto es el algoritmo de Viola-Jones que es un método de detección de objetos que destaca por su bajo coste computacional por lo que puede ser usado en sistemas de tiempo real y en general con objetos que contengan patrones típicos de iluminación [16].

Actualmente se usan sistemas biométricos como método para identificación de sujetos haciendo uso principalmente de la huella dactilar y el iris, con tasas de reconocimiento muy confiables, sin embargo, tienen como inconveniente la forma como se ingresa el registro al software, que para algunas aplicaciones pueden ser muy intrusivos, en los sistemas que usan reconocimiento facial se tiene la ventaja de que son de fácil adquisición por medio de cámaras que no interactúan directamente con los usuarios y que pueden estar en cualquier dispositivo móvil [5].

Existen diversos campos de aplicación para los sistemas de reconocimiento facial y actualmente son muy usados como parte del desarrollo de las ciudades inteligentes, en la figura 1 se puede observar varias áreas en los que se usan estos sistemas con sus respectivas aplicaciones.

Areas	Aplicaciones específicas
Entretenimiento	Videojuegos, realidad virtual, Interacción Humano-robot y humano-computadora.
Tarjetas Inteligentes	Licencias de conducir, ID's, pasaportes, inmigración.
Seguridad	Inicio de sesión o desbloqueo en dispositivos personales, encriptamiento de archivos, bases de datos.
Cumplimiento de la ley y vigilancia	Video-vigilancia avanzada, control de circuitos cerrados, control de accesos.

Figura 1. Áreas y aplicaciones de los sistemas de reconocimiento facial [17]

Es importante tener en cuenta que un clasificador con una precisión ideal solo puede llegar a coincidir con la de los humanos, por lo cual se requiere iluminación, expresión y calidad de imagen como factores para una adecuada detección [18].

Actualmente existen múltiples aplicaciones desarrolladas en el campo de reconocimiento facial, “*FACEFIRST*” [19] es usado por la fuerza pública para identificar sujetos y evitar falsos arrestos, permite hacer transacciones de forma segura, se usa como método de identificación en lugares de trabajo, casinos y en general lo adaptan a cualquier sistema.

“*FACE2GENE*” se usa en el área de la salud como herramienta de diagnóstico comparando con bases de datos de personas con trastornos que se ven reflejados en el rostro, utilizan algoritmos de aprendizaje profundo y clasificadores matemáticos que cuantifican las similitudes morfológicas [20].

Existe una aplicación reconocida por bancos como método de seguridad para identificar a los clientes, se llama “*FACEPHI*” en la cual se inscribe por medio de un formulario de registro que permite a los administradores introducir los datos del usuario y llevar a cabo la extracción de características faciales que compondrán el patrón facial del usuario [21].

Las anteriores aplicaciones son solo algunas de las decenas que existen actualmente en el mercado y que son utilizadas en áreas de la salud, investigación, industrias y seguridad. Los sistemas están muy desarrollados y se ha demostrado su funcionamiento mediante la aplicación práctica, el reto consiste entonces en desarrollar herramientas que permitan usar el reconocimiento facial en áreas que aún no ha sido usado, a pesar de los múltiples sistemas y aplicaciones desarrollados en este campo, no hay uno en el área de la salud que se use directamente para la identificación de los usuarios, el cual puede ser útil en casos donde no existan otras formas de identificar los pacientes.

### 3.4 Redes neuronales

Las redes neuronales son modelos que intentan simular el funcionamiento del sistema nervioso central, funciona por medio de unidades de procesamiento interconectadas que se parecen a versiones abstractas de las neuronas, están formadas por tres partes: la capa de entrada compuesta por la información que ingresa, una o varias capas ocultas compuestas por

las unidades de procesamiento y la capa de salida que representa el campo de destino de la propagación de la información [22].

En la figura 2 se puede observar el sistema global del proceso de una red neuronal, en la primera parte está el modelo de la **neurona** que consiste en varias entradas que representan las dendritas, un cuerpo de procesamiento y una salida, después se observa la **capa** formada por varias neuronas, en el tercer cuadro está la **red** que la conforman varias capas y en la última parte el sistema completo representado en función de la red neuronal, el algoritmo, una entrada y una salida [23].

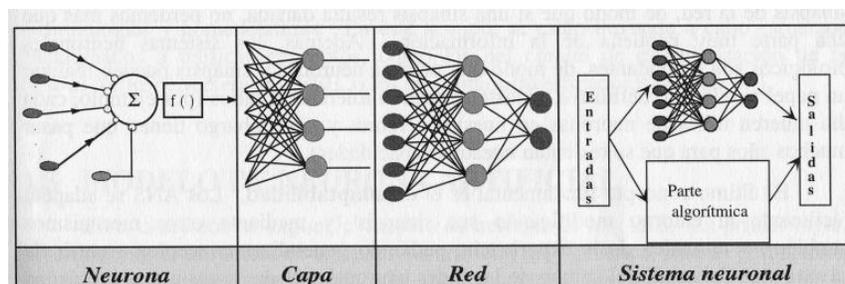


Figura 2. Sistema global del proceso de una red neuronal [23].

Una de las características más importantes de las redes neuronales es el **aprendizaje adaptativo** que consiste en el proceso de entrenamiento de la red por medio de ejemplos mediante los cuales aprenden a diferenciar patrones y el proceso termina siendo adaptable debido a la capacidad de autoajuste de las neuronas que componen el sistema [24].

### 3.4.1 Redes neuronales convolucionales

Las redes neuronales convolucionales suelen ser usadas para procesar imágenes, las tareas más comunes que se realizan son detección y categorización de objetos, clasificación de escenas y clasificación de imágenes en general. El procesamiento se basa en la operación de convolución donde cada pixel de salida es una combinación lineal de los pixeles de entrada [25].

La convolución es una operación que en tiempo continuo se define como la integral del producto de dos funciones después de desplazar una de ellas una distancia  $t$  [26]. Esta operación se puede aplicar sobre matrices y se utiliza principalmente como método de filtro para crear máscaras, contornos y en general ubicar elementos de interés en una imagen [27].

Las redes neuronales convolucionales presentan una arquitectura multicapa, donde cada capa está constituida por un número determinado de convoluciones con funciones de activación lineal, el tipo de conexión entre las neuronas está inspirado en la corteza visual de los animales, en la cual la respuesta de cada neurona puede presentarse matemáticamente como una operación de convolución [28], En la figura 3 se puede observar la estructura general de una red neuronal convolucional.

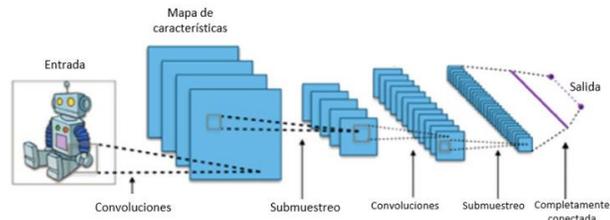


Figura 3. Estructura general de una red neuronal convolucional [27].

### 3.5 Bases de datos

Las bases de datos son un elemento fundamental en el entorno informático hoy en día, tienen aplicación práctica en muchos campos. Son una forma de gestionar datos y manejar información de forma muy práctica y organizada, además presenta muchas ventajas como el acceso eficiente para análisis, indexación y gestión [29].

MySQL que es un gestor de bases de datos relacionales de código abierto y con altos estándares de seguridad, dispone de una versión libre y otra licenciada, por lo cual es una excelente opción para realizar diseños y modelos usando la versión libre y en caso de escalar un proyecto a nivel comercial se puede usar la versión licenciada. [30]

### 3.6 JavaScript y TensorFlow

JavaScript es un lenguaje de programación que permite realizar actividades complejas y en especial, se usa en el desarrollo de páginas web, es ampliamente usado por la facilidad para interactuar con mapas, animaciones gráficas tridimensionales y en general interfaces de programación para todo tipo de aplicaciones [31].

Este lenguaje fue creado por Netscape y se usa principalmente para actuar en documentos HTML y darle el carácter dinámico que no puede conseguir por sí solo, originalmente fue implementado como parte de los navegadores web para que los scripts pudieran ser ejecutados y poder interactuar con los usuarios [32].

TensorFlow es una librería de código abierto que se usa en aplicaciones de aprendizaje automático, tiene muchas bibliotecas y recursos que son usados para desarrollo, además de modelos, entrenamiento y arquitectura para aplicaciones en código [33].

#### 3.6.1 Faceapi

Face-api es un módulo de JavaScript construido sobre tensorflow y funciona basado en redes neuronales convoluciones para detectar rostros en una imagen y ubicar los puntos en los contornos, optimizado para servicios web y móviles.

La aplicación se compone de tres etapas: la primera es la **detección del rostro** que la realiza utilizando el detector **SSD** (Single Shot Multibox detector) basado en la red **MobileNetV1**, agregando un algoritmo que realiza la predicción de la ubicación de los rostros en la imagen utilizando capas apiladas en la parte superior de la red. La segunda etapa ubica los **puntos de referencia faciales** a través de una red neuronal convolucional simple que entrega un vector con 68 características. La tercera etapa es la de reconocimiento facial que está basada en la arquitectura **ResNet-34** y está entrenada con un descriptor para aprender a mapear las características del rostro humano a partir de la imagen y los puntos de referencia faciales. Posteriormente usa la distancia euclidiana entre los descriptores de dos rostros y establece un valor umbral para determinar si se trata del mismo sujeto [34].

En la figura 4 se puede observar un modelo de la arquitectura de la red ResNet 34, consiste en un paso de convolución y agrupación seguida de 4 capas con un mismo patrón que es una convolución de 3x3 con una dimensión de mapa de características fijas [64, 128, 256 y 512] respectivamente, además las dimensiones de largo y ancho permanecen constantes sobre cada capa [35].

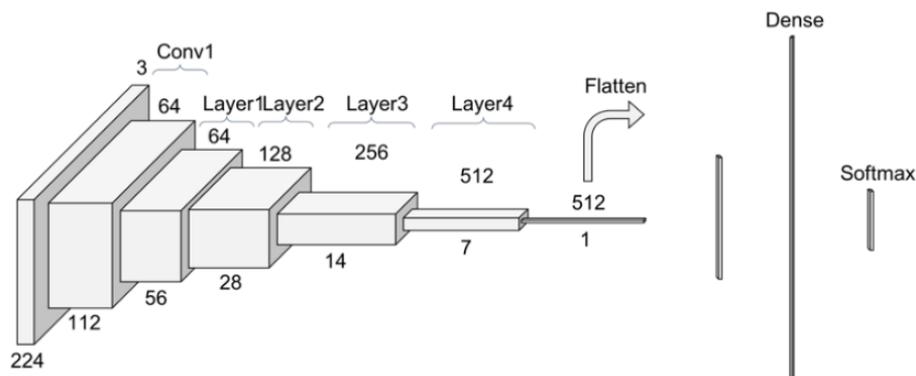


Figura 4. Arquitectura simplificada de ResNet 34 [35].

Existe gran variedad de detectores para determinar los puntos de referencia en el rostro. Sin embargo, la mayoría de métodos tratan de detectar las cejas, ojos, nariz, boca y mandíbula. Un método muy usado es el de Kazemi y Sullivan [36] que consiste en un conjunto de entrenamiento de puntos de referencia faciales etiquetados en una imagen (manualmente) especificando sus coordenadas (figura 5) y la probabilidad de distancia entre un par de píxeles de entrada, con esos datos se entrena un conjunto de árboles de regresión para extraer los puntos de referencia a partir de las intensidades de los píxeles.

No existe una distancia establecida para el valor umbral a partir del cual se puede considerar si se trata del mismo rostro. Sin embargo, se pueden realizar medidas y pruebas para establecer la precisión, la arquitectura **ResNet34** está implementada en dlib que tiene una

precisión de **99.38%** usando un valor de distancia de **0.6** [38]. En la figura 5 se puede observar la distribución de los puntos enumerados.

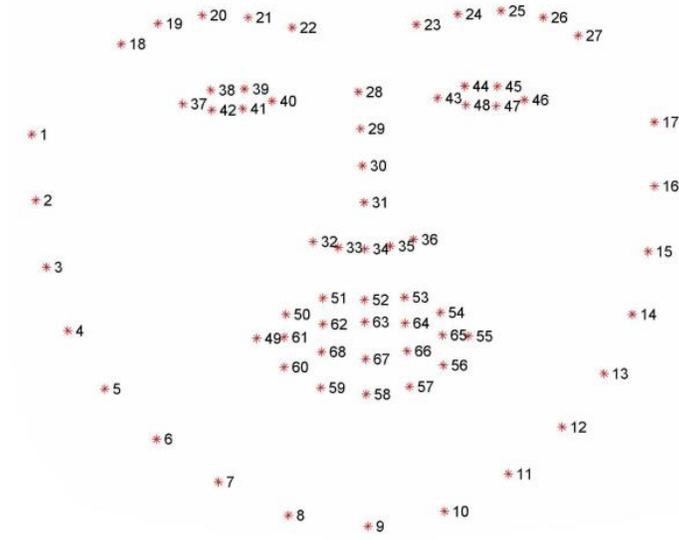


Figura 5. Distribución de 68 puntos de referencia en el rostro [37].

Faceapi tiene algunos modelos entrenados que son importados directamente de tensorflow, “*FaceLandmark68Net.ts*” [39] ubica los puntos de referencia en el rostro y a partir de esos puntos trabajan el modelo para estimar la edad y el género “*AgeGenderNet.ts*” [40] que está basado en las redes con clasificadores tf.Tensor1D para la edad y tf.Tensor2D para el género. Finalmente está “*FaceExpressionNet.ts*” [41] que reconoce 7 expresiones faciales: neutralidad, felicidad, tristeza, enojo, miedo, disgusto y sorpresa. Entrega un vector con las probabilidades para cada expresión asociadas a un rostro en una imagen.

## 4. Metodología

En la figura 6 se puede observar un diagrama de flujo con las actividades metodológicas que se realizaron para alcanzar los objetivos del proyecto, se diseñó de tal manera que se visualizaran las actividades en serie y en paralelo siguiendo una línea temporal de izquierda a derecha, la revisión bibliográfica fue una actividad transversal al desarrollo del proyecto.



Figura 6. Diagrama metodológico con actividades en serie y paralelo

### 4.1 Revisión bibliográfica

Para esta primera actividad se realizó una revisión del estado del arte de los algoritmos de reconocimiento facial, en donde se encontró información sobre los desarrollos que se han hecho alrededor de este tema y los métodos usados para hacer la detección de rostros a partir de una imagen, los resultados de esta actividad se pueden evidenciar en el marco teórico del proyecto.

### 4.2 Búsqueda de librerías para detección de rostros en JavaScript y Tensorflow

Al realizar la búsqueda bibliográfica se encontró que la mayor parte de la documentación de algoritmos y redes neuronales entrenadas para la detección de rostros estaba disponible para Python, sin embargo, se requería que el desarrollo del sistema se hiciera en un entorno web, por lo que la mejor opción era usar JavaScript y se encontró *face-api* que contenía los modelos para la detección de rostros y expresiones faciales. La descripción del funcionamiento se puede apreciar en el marco teórico.

### 4.3 Desarrollo de la interfaz gráfica de usuario

Para el diseño de la interfaz se disponía del acompañamiento de parte de ingenieros de sistemas y técnicos con experiencia en desarrollo, páginas web y aplicaciones; quienes fueron los encargados de brindar capacitación en el manejo de JavaScript para el diseño de la **interfaz gráfica, manejo de variables, librerías y lenguaje de etiquetado.**

La idea de usar JavaScript era montar el sistema en un servidor que se conectaba a una base de datos de tal manera que se pudiese acceder desde cualquier parte y facilitara su uso; Para esto se trabajó usando XAMPP que es un software de uso libre que permite trabajar en el servidor Apache que es de código abierto y puede interpretar los lenguajes Script, PHP y Perl.

#### **4.4 Diseño de la base de datos**

Teniendo en cuenta que el sistema se iba a montar sobre un servidor, se requería una conexión a una base de datos para mejorar la velocidad de procesamiento y tener buena capacidad de almacenamiento debido a la gran cantidad de datos que implicaba procesar imágenes y vídeos.

Durante el desarrollo de la base de datos, lo primero que se hizo fue investigar cual debería ser la información que contendría el registro, en este proceso se habló con algunos estudiantes y docentes de medicina de la universidad de Antioquia, se les explicó en qué consistía el proyecto que se estaba desarrollando y se pidió sus opiniones al respecto.

En general advirtieron que el tratamiento de datos era un tema muy delicado y requería ampliar la investigación para saber cuáles serían los datos que podrían asociarse a un sistema de reconocimiento, la recomendación general fue hacer genérico el registro de datos básicos como los contenidos en el documento de identidad que son los nombres, apellidos, Rh, edad y género; Además se sugirió incluir alergias, enfermedades y medicamentos en uso, que pueden ser información vital en una situación de emergencia, recomendaron primero validar bien el funcionamiento del software y al presentar el proyecto a las áreas de la salud, hacer la investigación formal de los datos que debería contener el registro.

Se creó la base de datos usando MySQL que es un gestor de bases de datos relacionales de código abierto y con altos estándares de seguridad, una opción práctica debido a que podía trabajar con la versión libre y en caso de escalar el proyecto la empresa disponía de una versión licenciada.

#### **4.5 Validación de modelos para identificación expresiones faciales, edad y género.**

El entrenamiento adecuado de una red neuronal requiere de miles de datos y existen varios modelos que pueden ser usados para el desarrollo de sistemas y aplicativos sin tener que desarrollar todos los algoritmos y entrenamientos desde el principio, es por eso que se investigó más sobre el módulo *face-api* y se encontró que tenía la librería *FaceExpressionNet.ts* con modelos de Tensorflow entrenados para la estimación de 7 expresiones faciales: neutralidad, tristeza, felicidad, miedo, disgusto, enojo y sorpresa.

Al observar todas las herramientas que podían usarse, se decidió usar la librería *AgeGenderNet.ts* para realizar pruebas sobre la estimación de la edad y el género para después establecer su exactitud y viabilidad en un sistema para áreas de la salud.

Se hicieron pruebas preliminares de funcionamiento y se tenía acceso al nivel de confianza que es un valor entre 0 y 1 que cuantifican la probabilidad distribuida de que un rostro esté realizando alguna de esas expresiones, además entregaba el valor de probabilidad de que se tratara del rostro de un hombre o de una mujer, y en función de esos parámetros se podía mostrar en la pantalla el resultado de las estimaciones.

#### 4.6 Diseño del sistema de reconocimiento en tiempo real

Para tener una visión general del funcionamiento del proceso de reconocimiento se diseñó un esquema de la arquitectura que detalla cómo se deberían unir y comunicar cada uno de los componentes del sistema.

En la figura 7 se puede observar el esquema, las flechas que indican las direcciones en que se dan las interacciones del sistema, la interacción entre el usuario y la interfaz es bidireccional, el usuario ingresa los datos de un sujeto y una imagen de registro, luego el sistema retorna una detección con los parámetros estimados (Edad, género, expresión facial y nombre en caso de estar registrado), para hacer esto los datos que ingresan por la interfaz son enviados a la aplicación que a su vez los reenvía a la base de datos para almacenarlos en caso de un registro y para compararlos en caso de una búsqueda de sujeto, finalmente el resultado de la comparación se retorna nuevamente a la interfaz para ser visualizado por el usuario.

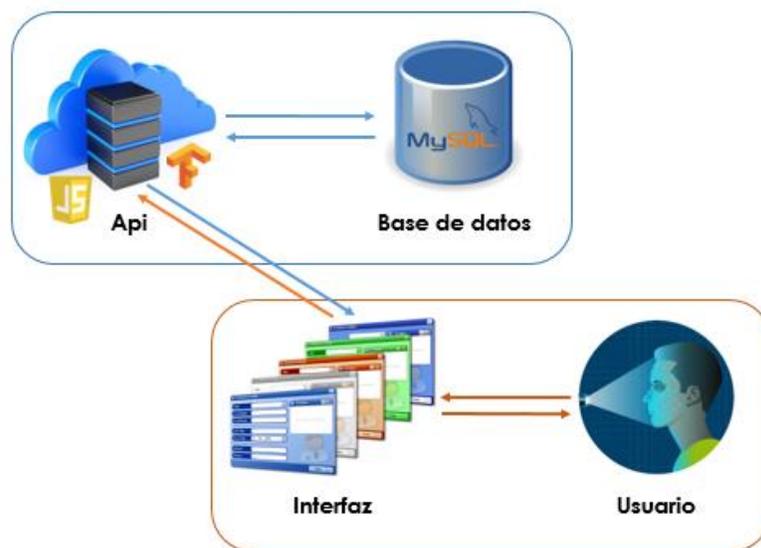


Figura 7. Arquitectura general del proceso

Después de diseñar la arquitectura del sistema, se desarrollaron los algoritmos para la interacción de los componentes y se realizaron pruebas para determinar la estabilidad del sistema al realizar la detección en tiempo real por medio de grabación de vídeo.

#### 4.7 Registros de datos biométricos de rostros con su respectiva información

Teniendo el sistema funcionando y habiendo realizado varias pruebas preliminares para determinar su funcionalidad, se procedió a realizar pruebas estructuradas para validar todos los modelos y datos que devuelven los algoritmos.

Se diseñó una prueba en la cual se establecía realizar un registro a 50 sujetos, cada uno consistía en tomar una fotografía del rostro del sujeto y registrar su información que posteriormente quedaba almacenada en la base de datos.

Con el fin de probar el reconocimiento en condiciones que no se apreciaran todos los rasgos del rostro, se hicieron modificaciones a las fotos del registro añadiendo accesorios: gafas transparentes, gafas oscuras, peluca, gorra, y rostro hinchado. En la figura 8 se pueden observar los registros con diferentes modificaciones. A cada uno de los 50 sujetos se les realizó todas las modificaciones, para efectos prácticos se muestran 10 sujetos por cada modificación.

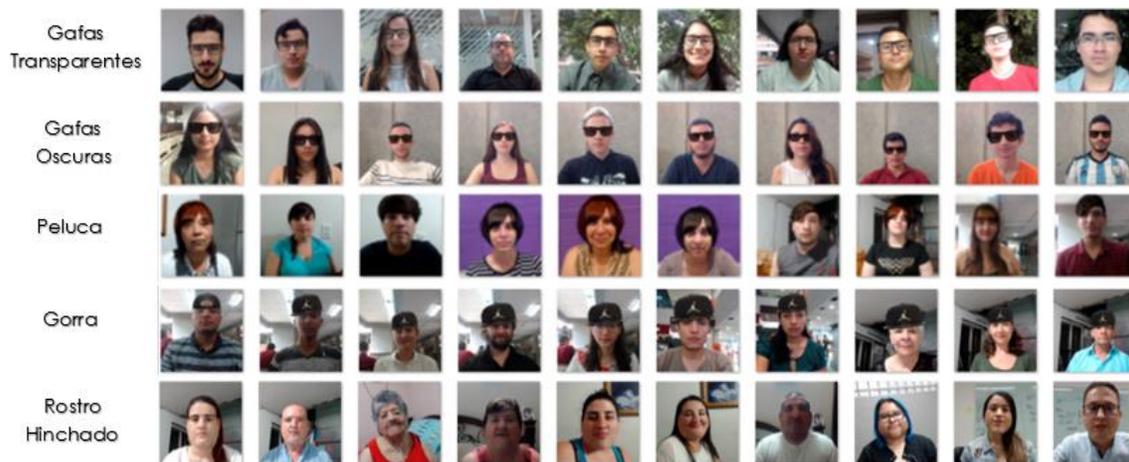


Figura 8. Modificaciones aplicadas a los registros.

#### 4.8 Diseño de pruebas de desempeño

Esta actividad se planteó para evaluar el sistema desarrollado y determinar posibles mejoras a partir de la percepción de las personas al interactuar con la interfaz y los distintos elementos de la aplicación web.

Para el desarrollo de la prueba se le solicitó a cada sujeto que usara la aplicación web de la siguiente manera: primero hacía el registro de su propio rostro con los respectivos datos y posteriormente realizaba la búsqueda en la base de datos por medio de fotos que podía tomar con una cámara que se le proporcionaba, después se le pedía adjuntar archivos para que realizara diversas búsquedas de personas que ya estaban registradas y finalmente se les pedía que usaran la opción de grabar vídeo para que vieran el reconocimiento en tiempo real y trataran de simular diferentes expresiones faciales para observar los cambios en los parámetros de detección.

Al finalizar la prueba, a cada sujeto se le pedía responder una encuesta para calificar la aplicación desde el punto de vista de **funcionamiento (F)**, **diseño (D)** y **usabilidad (U)**. A cada pregunta se le podía asignar una calificación entre 0 y 5 marcando con una X el valor correspondiente. El formato de la encuesta se puede observar en la figura 9.

Prueba de desempeño para sistema de reconocimiento facial		Calificación				
La siguiente prueba tiene como objetivo evaluar el sistema de reconocimiento a partir de algunas preguntas sobre funcionamiento (F), diseño (D) y usabilidad (U). Se deben responder asignando una calificación numerica de 0 a 5, donde 0 es insatisfecho y 5 es muy satisfecho.						
F1	¿Es posible registrar el rostro de un sujeto y su información personal?					
F2	¿Se puede adjuntar una imagen y realizar un registro?					
F3	¿Se puede tomar una foto y reconocer un sujeto a partir de esta?					
F4	¿A partir de una imagen adjunta es posible realizar el reconocimiento y visualizar la información previamente registrada?					
F5	¿Se visualiza en tiempo real un rectángulo que sigue el rostro, muestra información de género, edad, expresión facial y el nombre en caso de estar registrado?					
D1	¿Es estética la interfaz?					
D2	¿El diseño facilita el manejo del sistema?					
D3	¿Es fluido el reconocimiento en tiempo real?					
U1	¿Es fácil diferenciar entre registro de sujeto y búsqueda de sujeto?					
U2	¿Se puede cambiar entre tomar foto y adjuntar archivo con facilidad?					
U3	¿Se activa el sistema en modo reconocimiento en tiempo real con facilidad?					
U4	¿El ingreso de datos al sistema y el registro se hace de forma intuitiva y sencilla?					
Nombre:		Ocupación:				
Comentarios:						

Figura 9. Encuesta para evaluar desempeño

#### 4.9 Diseño de pruebas para evaluar identificación de expresiones faciales, edad y género.

Estas pruebas fueron realizadas con el fin de cuantificar la precisión de los parámetros que estimaba el sistema cuando detectaba el rostro, para las **expresiones faciales** se le pidió a cada uno de los sujetos que componen el registro que tratara de simular las 7 expresiones faciales que detectaba el sistema que eran neutralidad, tristeza, felicidad, miedo, disgusto, enojo y sorpresa. La estructura de la tabla donde se registraron los resultados, se puede observar en la figura 10.

Registro de resultados para expresiones faciales							
Sujeto	Expresiones faciales						
	Neutral	Feliz	Triste	Enojado	Miedo	Sorprendido	Disgustado
S1							
S2							
S3							
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
S50							

Nota: Se marca con una "x" la casilla cuando el software logra reconocer la emoción simulada por el sujeto

Figura 10. Tabla para el registro de los resultados de reconocimiento de expresiones faciales

## 5. Resultados

### 5.1 Detección

Utilizando el módulo “*face-api.js*” y “*tensorflow/tfjs*”, se desarrolló un código que permitía adjuntar una imagen y devolver un rectángulo que encerraba el rostro detectado por medio de las redes neuronales contenidas en los módulos.

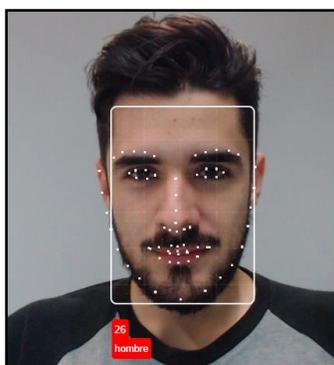
En la figura 11 se puede observar el resultado de la fase de **detección** en la cual el algoritmo detectaba el rostro en la imagen y lo encerraba en un rectángulo de color blanco para la interfaz de usuario.



*Figura 11. Etapa de detección de rostros*

### 5.2 Puntos de referencia faciales, edad y género

Después de detectar el rostro en la imagen, seguía la etapa en la que se ponían los **puntos de referencia** en el rostro y se estimaban los parámetros de edad y género, en la figura 12 se pueden observar los puntos sobre el rostro y los parámetros en la parte inferior izquierda del recuadro que encierra el rostro.

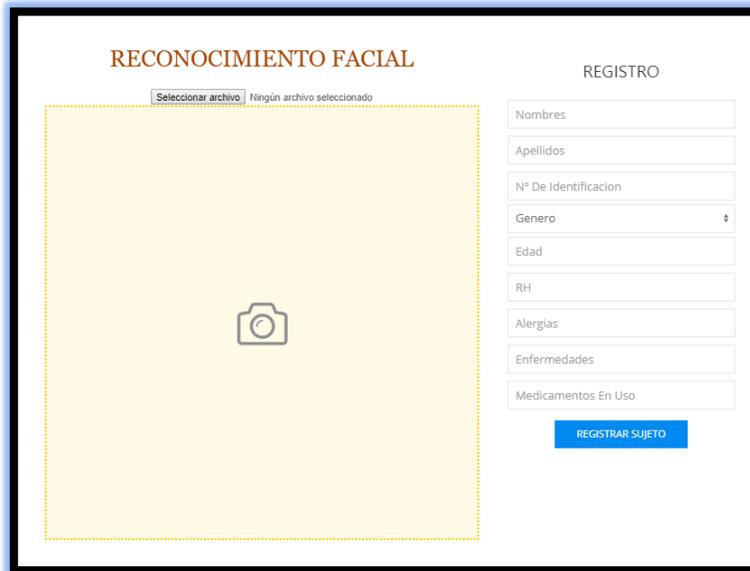


*Figura 12. Puntos de referencia y estimación de edad y género.*

### 5.3 Interfaz gráfica

Se desarrolló la interfaz de registro utilizando un servidor local, en la figura 13 se puede observar el diseño que contiene en la parte superior izquierda el título y las opciones para tomar una foto o adjuntar una imagen; En la parte derecha se pueden ver campos para

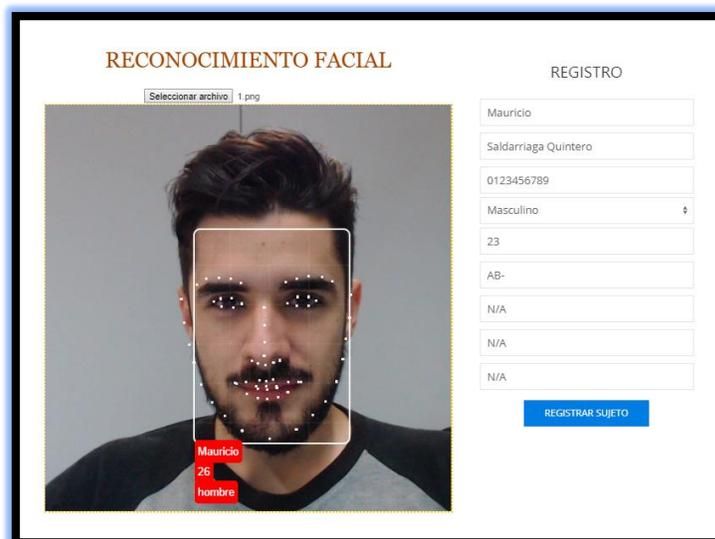
completar la información con los parámetros en gris claro que indica la información que se debe registrar y al final un botón de registro que completa el proceso guardando los datos.



The screenshot shows a web interface for facial recognition registration. On the left, there is a large yellow dashed box containing a camera icon and a file selection button labeled 'Seleccionar archivo' with the text 'Ningún archivo seleccionado' below it. On the right, there is a 'REGISTRO' section with several input fields: 'Nombres', 'Apellidos', 'N° De Identificación', 'Genero' (with a dropdown arrow), 'Edad', 'RH', 'Alergias', 'Enfermedades', and 'Medicamentos En Uso'. Below these fields is a blue button labeled 'REGISTRAR SUJETO'.

Figura 13. Interfaz para el registro de sujetos.

Después de presionar el botón de registro aparecía el recuadro encerrando el rostro del sujeto con los puntos marcados y la estimación de edad y género, además se agregaba el nombre de registro de la persona en la parte inferior izquierda y esa era la confirmación de que el registro se había realizado de forma exitosa con el registro había sido almacenado en la base de datos, este proceso se muestra en la figura 14.



The screenshot shows the same web interface as Figure 13, but now with a photo of a man's face. A white bounding box is drawn around the face, and a red box in the bottom left corner of the photo contains the text 'Mauricio', '26', and 'hombre'. The 'REGISTRO' form on the right is now filled with the following information: 'Mauricio', 'Saldarriaga Quintero', '0123456789', 'Masculino' (with a dropdown arrow), '23', 'AB-', 'N/A', 'N/A', and 'N/A'. The blue 'REGISTRAR SUJETO' button is still present.

Figura 14. Interfaz al finalizar el proceso de registro.

La imagen quedaba guardada como un vector que contenía las características determinadas por el descriptor facial, luego estas eran usadas para hacer el reconocimiento cuando se ingresara una imagen y se hiciera la comparación con los descriptores de todos los sujetos registrados.

#### 5.4 Base de datos

El diseño de la base de datos se realizó usando como variables la información que se observa en la interfaz de la figura 13, adicionalmente se incluyó la fecha en la cual se realizaba cada registro y el vector que contenía las características del rostro.

En la figura 15 se puede observar la interfaz de MySQL y como se estructuraron los datos teniendo en cuenta el tipo de variables, el primer dato es el “*id*” que es el parámetro principal y es el que contiene el conteo de los registros que van ingresando a la base de datos, entre las filas 2 y 10 se encuentran las variables del registro definidas como “*varchar*” (variable tipo carácter) y entre paréntesis se encuentran sus longitudes máximas que se pusieron de forma que no hubiese conflicto con la cantidad de datos que podría llegar a contener cada campo, la variable “*photo*” es de tipo texto y contiene el vector de características de la imagen de registro y finalmente la fecha que es un valor que se autocompleta al recibir un registro.

#	Nombre	Tipo	Cotejamiento	Atributos	Nulo	Predeterminado	Comentarios	Extra	Acción
1	id	int(11)			No	Ninguna		AUTO_INCREMENT	Cambiar Eliminar Más
2	nombre	varchar(100)	latin1_swedish_ci		No	Ninguna			Cambiar Eliminar Más
3	apellido	varchar(100)	latin1_swedish_ci		No	Ninguna			Cambiar Eliminar Más
4	identificacion	varchar(100)	latin1_swedish_ci		No	Ninguna			Cambiar Eliminar Más
5	genero	varchar(20)	latin1_swedish_ci		No	Ninguna			Cambiar Eliminar Más
6	edad	varchar(20)	latin1_swedish_ci		No	Ninguna			Cambiar Eliminar Más
7	rh	varchar(20)	latin1_swedish_ci		No	Ninguna			Cambiar Eliminar Más
8	alergias	varchar(100)	latin1_swedish_ci		No	Ninguna			Cambiar Eliminar Más
9	enfermedades	varchar(100)	latin1_swedish_ci		No	Ninguna			Cambiar Eliminar Más
10	medicamentos	varchar(100)	latin1_swedish_ci		No	Ninguna			Cambiar Eliminar Más
11	photo	text	latin1_swedish_ci		No	Ninguna			Cambiar Eliminar Más
12	fecha	date			No	current_timestamp()			Cambiar Eliminar Más

Figura 15. Interfaz de MySQL con estructura de datos de registro.

Después se realizaron algunos registros para visualizar la información en la base de datos que se puede observar en la figura 16, para efectos de pruebas se podían poner datos que no fuesen verdaderos en los campos de registro, ya que no era objetivo del proyecto realizar registros reales, sino evaluar la funcionalidad del sistema.

	id	nombre	apellido	identificacion	genero	edad	rh	alergias	enfermedades	medicamentos	photo	fecha
Editar Copiar Borrar	1	Mauricio	Saldarriaga Quintero	1000123456	Masculino	23	AB-	Polvo	Ansiedad	Amtriptilina	["0"-0.14592452347278595,"1"-0.12164703011512756,...	2019-11-18
Editar Copiar Borrar	2	Alejandro	Arias Rodríguez	1000123457	Masculino	25	A+	Mariscos	N/A	N/A	["0"-0.14993919918323242,"1"-0.04854122890419008...	2019-11-18
Editar Copiar Borrar	3	Karen	Calle Betancur	1000123458	Femenino	24	O+	Flores	N/A	N/A	["0"-0.057389892937729945,"1"-0.04117397218942642...	2019-11-18
Editar Copiar Borrar	4	Juan	Villada	1000123459	Masculino	35	A+	N/A	Tinitus	Mareol	["0"-0.06386071850081417,"1"-0.07181112468242645...	2019-11-18

Figura 16. Visualización de datos de registro en MySQL.

## 5.5 Expresiones faciales

En la figura 17 se puede visualizar los resultados que arroja el modelo de expresiones al ingresar una imagen que contiene un rostro, si se suman todas las probabilidades asociadas a cada expresión obtendríamos un valor de 1, en este caso si tomamos el valor como un porcentaje se vería un peso de 98% sobre la expresión de neutralidad y es fácil inferir que los valores que toman las demás expresiones son insignificantes.



Figura 17. Probabilidades para la estimación de la expresión facial.

Se le añadió en la interfaz la visualización de la expresión facial en la parte superior izquierda del rectángulo que encierra el rostro detectado y para efectos de pruebas se trató de simular las 7 expresiones faciales y los resultados se muestra en la figura 18.

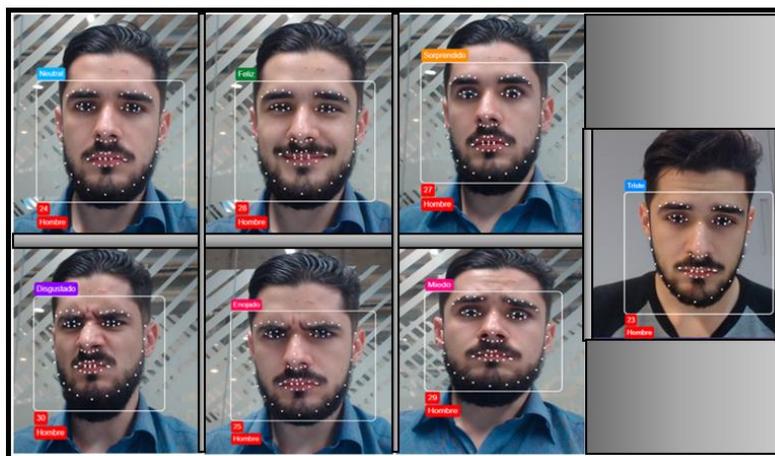


Figura 18. Simulación de expresiones faciales, la estimación se muestra en la parte superior izquierda del rectángulo que encierra el rostro.

## 5.6 Reconocimiento en tiempo real

El sistema devolvía en tiempo real la información del nombre de un sujeto previamente registrado, mostraba todo el tiempo variaciones en los parámetros de género, edad y expresión facial; En caso de no estar registrado seguía mostrando todos los parámetros, pero en el lugar que debería estar el nombre aparecía “Desconocido”, esta información se evidencia realizando una captura de pantalla durante la grabación de un vídeo con un sujeto registrado y otro que no lo está, esto se muestra en la figura 19.

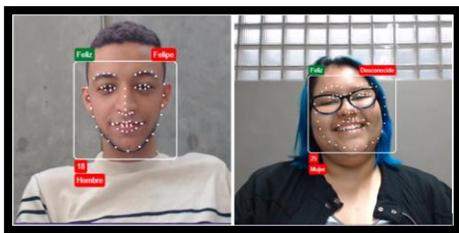


Figura 19. Reconocimiento en tiempo real, a la izquierda un sujeto registrado y a la derecha un sujeto sin registrar.

## 5.7 Pruebas de reconocimiento editando los registros

Se editaron los registros agregando gafas transparentes y oscuras a los rostros para intentar ocultar los rasgos de los ojos y las cejas, teniendo un mayor efecto para el reconocimiento el uso de gafas oscuras. La idea de la gorra era cubrir parcial o totalmente las cejas, el uso de la peluca cubría rasgos parcialmente en un lado de la cara y el rostro hinchado tenía el efecto de separar los puntos de detección y simulaba una reacción alérgica. En la tabla 1 se puede observar los resultados de estas pruebas.

Tabla 1. Resultados estadísticos del porcentaje de reconocimiento al variar características en 50 rostros

Resultados de reconocimiento al modificar registros					
Parámetros	Gafas transparentes	Gafas Oscuras	Peluca	Gorra	Cara hinchada
$\bar{x}$	67%	65%	68%	65%	62%
$\sigma$	11%	6%	7%	7%	8%
Cv	16%	9%	11%	10%	12%
Reconocidos	43	42	43	39	34
Porcentaje reconocidos	84%	82%	84%	76%	67%

La variable “porcentaje de similitud” es a la que hace referencia las medidas de tendencia central mostradas en la tabla 1, que son el promedio ( $\mu$ ), la desviación estándar ( $\sigma$ ) y el coeficiente de variación (Cv), teniendo en cuenta que la población de estudio fueron los 50 rostros registrados y que la cantidad de reconocidos se obtenía contando el número de sujetos que superaran un valor de similitud mayor al 60%.

## 5.8 Pruebas de desempeño

Esta prueba se le aplicó a un total de 20 personas, se buscaba que la muestra fuese variada en cuanto a conocimientos técnicos de las personas que respondían la encuesta mostrada en la figura 9, por lo que algunas se realizaron en la universidad de Antioquia a estudiantes de distintas ingenierías y se acudió al personal administrativo para que no tuviesen conocimientos muy técnicos en manejo de software.

Los resultados de las calificaciones de las preguntas sobre funcionalidad están resaltados con color azul, diseño color verde y usabilidad color amarillo. Los resultados de las encuestas se pueden observar en la tabla 2 que fue diseñada de tal manera que se observen las calificaciones de cada pregunta y los promedios diferenciando cada sujeto, cada pregunta y cada grupo de preguntas. En la figura 20 se observa el diagrama de barras comparando cada sujeto con el promedio global de su calificación y en la figura 21 el diagrama comparando cada pregunta con su calificación promedio.

Tabla 2. Resultados de las encuestas en pruebas de desempeño

Resultados de las encuestas													
Sujeto/Pregunta	F1	F2	F3	F4	F5	D1	D2	D3	U1	U2	U3	U4	$\bar{x}$ sujeto
S1	5	5	5	5	3	3	3	4	3	5	5	5	4,25
S2	5	5	5	5	5	4	4	4	3	5	5	5	4,58
S3	5	5	5	5	5	2	3	3	1	5	3	5	3,92
S4	5	5	4	3	3	1	5	3	2	5	5	5	3,83
S5	5	5	5	4	5	2	3	4	1	4	5	5	4,00
S6	5	5	5	5	5	3	4	5	5	5	5	5	4,75
S7	5	5	5	5	5	4	5	3	5	5	4	5	4,67
S8	5	5	5	5	5	4	5	5	3	4	2	5	4,42
S9	5	5	5	5	5	4	4	5	4	4	5	4	4,58
S10	5	5	5	5	5	2	4	3	3	5	3	5	4,17
S11	5	5	4	4	4	1	3	4	2	3	0	4	3,25
S12	5	5	5	5	5	5	5	5	5	5	5	5	5,00
S13	3	5	4	5	5	3	2	4	4	1	5	3	3,67
S14	5	5	5	5	4	4	3	4	5	5	5	4	4,50
S15	5	5	5	5	5	4	4	4	4	4	5	5	4,58
S16	5	5	5	5	5	5	5	4	5	5	4	5	4,83
S17	5	5	5	5	5	5	5	5	5	3	5	5	4,83
S18	5	5	4	4	5	5	5	5	5	5	5	5	4,83
S19	5	5	5	5	5	4	5	5	1	4	3	5	4,33
S20	4	4	4	4	5	5	5	4	3	5	5	5	4,42
$\bar{x}$ pregunta	4,85	4,95	4,75	4,70	4,70	3,50	4,10	4,15	3,45	4,35	4,20	4,75	4,37
$\bar{x}$ grupos	4,79					3,92			4,19				

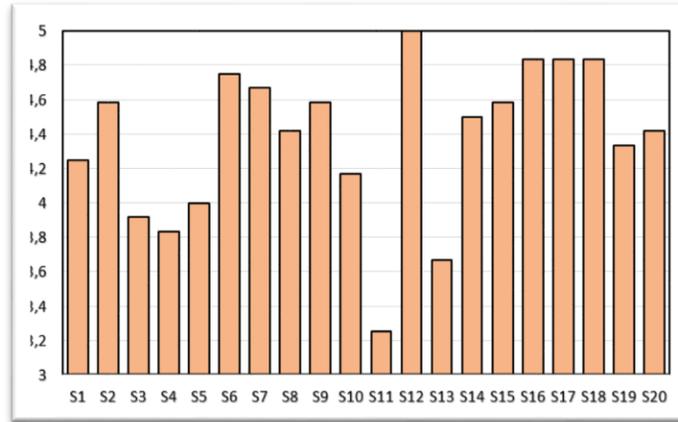


Figura 20. Calificación promedio para cada sujeto en la prueba de desempeño.

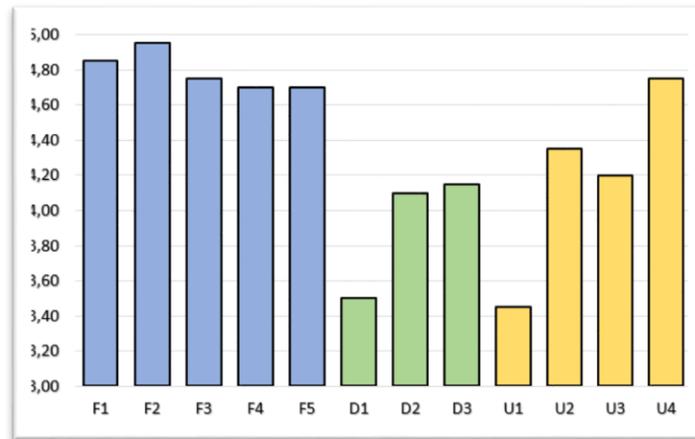


Figura 21. Calificación promedio para cada pregunta en la prueba de desempeño.

## 5.9 Validación de parámetros

En la tabla 3 se pueden ver los resultados de las pruebas realizadas para las expresiones faciales y resume los resultados de las pruebas que se registraron en la estructura que muestra la figura 10.

Tabla 3. Resumen de los resultados de las pruebas aplicadas para la estimación de expresiones faciales

Resultados de pruebas de expresiones faciales							
N/A	Neutral	Feliz	Triste	Enojado	Miedo	Sorprendido	Disgustado
Total	49/50	49/50	35/50	18/50	4/50	41/50	10/50
Porcentaje	98%	98%	70%	36%	8%	80%	20%

En la tabla 4 se muestra el resumen de los datos registrados sobre el género, en donde se comparaba el género que entregaba el sistema, con el real del sujeto registrado.

*Tabla 4. Resultados de reconocimiento de genero*

Identificación de genero			
N/A	Cantidad	Aciertos	Porcentaje
Masculino	24	24	100%
Femenino	26	23	88%
Total	50	47	94%

Para los resultados de la estimación de edad, se calculaba el porcentaje de error con respecto a la edad real que era un dato de registro y los resultados eran errores negativos para estimaciones por debajo de la edad real y positivos para valores estimados por encima, en el histograma de la figura 22 se puede observar el porcentaje de error para cada sujeto, diferenciando con colores, azul para porcentajes positivos y rojo para porcentajes negativos.



*Figura 22. Histograma de los porcentajes de error para la estimación de la edad, en rojo porcentajes negativos y en azul los positivos*

## **6. Análisis de resultados**

### **6.1 Interfaz gráfica y reconocimiento**

Se encontró una herramienta muy potente que es el módulo “*face-api*” el cual mostró en las pruebas preliminares que detectaba de forma correcta el rostro de una persona, ubicaba los puntos de referencia faciales y extraía las características para estimar las expresiones faciales, la edad y el género.

Se desarrolló una interfaz sencilla con el fin de que pudiese ser usada por cualquier usuario sin necesidad de conocimientos técnicos ya que el ingreso de datos se hacía de forma intuitiva y se disponía de dos opciones para el registro de la imagen: adjuntar archivo y tomar foto.

Se podía ver en la interfaz la expresión facial en la parte superior izquierda y a cada expresión se le asignaba un color diferente de la siguiente manera: azul claro para neutral, verde para feliz, naranja para sorprendido, morado para disgustado, rojo para enojado, rosado para miedo y azul oscuro para triste; Esto con el fin de hacer más interactiva la interfaz e identificar más fácil el cambio en las expresiones.

Se debe tener en cuenta que las expresiones faciales varían en cada persona y los modelos solo tratan de hacer una estimación basados en las distancias relativas de los puntos marcados en el rostro y las bases de datos con miles de imágenes utilizadas para el entrenamiento del algoritmo.

Para el sistema en tiempo real se observó durante las pruebas que el parámetro más estable era el género que permanecía constante a lo largo de las grabaciones, luego seguía la expresión facial que solo variaba cuando habían cambios bruscos en los movimientos de los músculos del rostro, después estaba la edad que variaba todo el tiempo durante el vídeo con cualquier movimiento del sujeto o cambio en su expresión facial, sin embargo, siempre oscilaba alrededor de valores similares para un mismo, el nombre del sujeto se podía observar en caso de estar registrado, este fue el parámetro con menor estabilidad debido a que el procesamiento exigía cada vez mayor gasto computacional a medida que crecía la cantidad de registros en la base de datos, se debía comparar todo el tiempo con todos los sujetos y los porcentajes de similitud se iban alternando con los cambios en las expresiones que los sujetos probaban durante el vídeo.

### **6.2 Base de datos**

Se diseñó una base de datos que tenía una comunicación efectiva con el sistema, al terminar de hacer un registro el tiempo que tomaba en llegar toda la información a la base de datos era muy corto y esto se podía comprobar sabiendo que al presionar el botón de registro lo primero que hacía el algoritmo era enviar la información a la base de datos y luego aparecía en la imagen un recuadro encerrando el rostro y marcando los puntos de detección, midiendo el tiempo entre esos dos eventos se podía saber cuánto tiempo tomaba el almacenamiento de

datos, esta información era mostrada en la consola y en todos los registros que se realizaron en el desarrollo del proyecto nunca mostró un tiempo mayor a 2 segundos.

### 6.3 Pruebas de reconocimiento al realizar modificaciones

Cada vez que se realizaba una búsqueda ingresando una foto, el sistema comparaba usando la distancia euclidiana entre las características extraídas de los puntos de referencia faciales con cada uno de los sujetos que habían registrados en la base de datos, devolviendo en la interfaz gráfica el valor que tuviese un porcentaje de similitud más alto. Sin embargo, detrás de la interfaz el algoritmo entregaba los porcentajes de reconocimiento asociado a cada sujeto registrado con el fin de saber que valores entregaba cuando no se trataba del sujeto que se estaba buscando. En los 50 registros que se realizaron ningún valor superó el 40% cuando no se trataba del sujeto que se estaba buscando, esto es muy importante debido a que es una base para establecer el parámetro de reconocimiento. que se definió en 60% para ser prudente con el reconocimiento, estando 20% por encima de cualquier resultado que pudiese entregar un reconocimiento falso positivo.

Teniendo en cuenta el análisis anterior se puede entender mejor los resultados mostrados en la tabla 1, se puede ver que todos los promedios calculados sobre cada modificación realizada en los rostros estuvieron por encima del valor umbral que establece un reconocimiento adecuado, teniendo en cuenta que el valor promedio es un dato que de por sí solo no dice mucho, se calcularon sus respectivas desviaciones estándar y coeficientes de variación, a nivel estadístico se dice que este coeficiente debe ser menor a 20% para que el promedio sea un valor que representa la muestra de estudio y en la tabla se puede observar que todos cumplen ese criterio.

Teniendo definidas las medidas de tendencia central, se debía hacer un conteo manual de los valores debido a que las desviaciones estándar indicaban que podían existir sujetos con valores de reconocimiento menor al 60%, y se encontró que la cantidad de sujetos reconocidos usando gafas transparentes y peluca fueron los de mayor éxito con un 84% correspondiente a 43 sujetos, un resultado esperado teniendo en cuenta que eran las modificaciones que menos efecto tenían sobre los rasgos faciales, le siguió los de gafas oscuras con un éxito del 82% correspondiente a 42 sujetos, el cual es un porcentaje bastante alto, realizando las pruebas se observó que el algoritmo ponía los puntos en los ojos y las cejas a pesar de que las gafas eran más oscuras, lo que hacía conservar las posiciones en los puntos de comparación. Después estaban los sujetos de gorra con un éxito del 76% correspondiente a 39 sujetos, un porcentaje menor pero que puede ser explicado debido a que está modificación ocultaba completamente las cejas y el software ponía los puntos de una forma menos precisa en esas regiones; Finalmente los rostros hinchados tuvieron el menor porcentaje de éxito con tan solo 67% correspondiente a 34 sujetos, esto puede explicarse debido al cambio de las posiciones en los puntos de detección.

En general los resultados estuvieron muy satisfactorios debido a que el sistema seguía teniendo gran cantidad de aciertos en el reconocimiento a pesar de las modificaciones que se realizaban a los rostros, **cabe aclarar que estas modificaciones se realizaron con objetos virtuales** y haría falta realizar las pruebas con objetos reales.

## 6.4 Pruebas de desempeño

En la tabla 2 se pueden observar las calificaciones realizadas para las preguntas que se plantean en la encuesta de la figura 9, en la última columna, leyendo la tabla de izquierda a derecha se puede ver el promedio de todas las calificaciones realizada por cada uno de los sujetos y está información se visualiza en el diagrama de barras de la figura 20.

Al realizar los cálculos se encontró que la calificación promedio de las pruebas de desempeño fue de 4,37 con una desviación estándar de 0,45, la relación entre esos dos parámetros se usó para calcular el coeficiente de variación, que en este caso fue de aproximadamente 10%, un valor relativamente bajo que indicaba que las opiniones respecto al sistema eran muy similares.

Se estableció un rango de calificaciones entre 3,92 y 4,82 tomando una desviación estándar por encima y una por debajo del promedio, los sujetos que dieron un promedio de calificación por debajo de 3,92 fueron los más exigentes y los que dieron un promedio de calificación por encima de 4,82 fueron los más conformistas.

El sujeto con la calificación más baja fue el 11, se trataba de una diseñadora gráfica que tuvo una opinión negativa con respecto al diseño de la aplicación, en los comentarios escribió sobre una interfaz simple y poca estética, respecto a las preguntas de usabilidad comentó que había varios elementos poco intuitivos como el botón que sirve para cambiar entre registro y búsqueda de sujetos (Ese comentario con respecto a la usabilidad fue realizado por varios usuarios).

El sujeto que dio la mayor calificación fue el 12, se trataba de una tecnóloga de salud y seguridad en el trabajo, durante la prueba estaba muy fascinada con el proyecto debido a la gran utilidad y viabilidad que le vio y su opinión pudo estar sesgada por ese mismo motivo, debido a que si había algunas falencias en la interfaz y la usabilidad.

Analizando los promedios por conjuntos de preguntas, los cuales se pueden observar en la última fila de la tabla 2 y se visualizan en el diagrama de barras de la figura 21, se encuentra que las preguntas sobre funcionalidad tuvieron el puntaje más alto con un valor de calificación promedio de 4,79, un excelente resultado ya que indicaba que el sistema funcionaba correctamente y hacía los reconocimientos de forma adecuada.

Las preguntas sobre usabilidad obtuvieron un promedio de calificación de 4,19 y esto se debe a lo que se mencionó anteriormente sobre la ubicación del botón que cambia entre registro y búsqueda, además se sugirió establecer el tipo de sangre como una lista desplegable teniendo en cuenta que las respuestas posibles son muy reducidas.

Las preguntas con el promedio más bajo de calificación fueron las que hacían referencia al diseño con un valor de 3,92 siendo una calificación aceptable. Sin embargo, en la recopilación de comentarios se encontró que la percepción de la estética de la interfaz fue regular, la pregunta D1 era la que se refería a la estética y tuvo uno de los promedios más bajos con 3,5 y fue superada solo por U1 con un valor de 3,45 debido a la forma de realizar el cambio entre registro y búsqueda.

## 6.5 Expresiones faciales

Los resultados de las pruebas para expresiones faciales se resumen en la tabla 3, es importante tener en cuenta que las expresiones faciales fueron forzadas y muchos sujetos durante las pruebas manifestaban no saber simular algunas de ellas, adicionalmente se confundían cuando se les pedía enojo y disgusto, ya que no diferenciaban entre ellas, sin embargo, bastaba con explicarles que el disgusto tenía que ver más con las reacciones que generan asco y que el enojo tiene más que ver con la sensación de molestia.

Las expresiones que reconoció el sistema con mayor facilidad fueron felicidad y neutralidad con un porcentaje de éxito del 98% correspondiente a 49 sujetos de 50, esto debido a que eran las más fáciles de diferenciar y simular por medio de gestos, la felicidad se evidenciaba fácilmente por medio de una sonrisa que en termino de los puntos de referencia se organizan en forma de una curva que abre hacia arriba y la neutralidad dejaba los músculos del rostro y los puntos en su posición normal.

Después le siguió sorpresa con un porcentaje de reconocimiento del 80% correspondiente a 41 sujetos, esta también es una expresión fácil de simular ya que se puede evidenciar con la apertura de la boca y los músculos alrededor de los ojos, lo cual generaba la separación de los puntos de detección en esas regiones, después estaba la tristeza con un éxito del 70% correspondiente a 35 sujetos, está es una expresión que varía mucho entre las personas. Sin embargo, al intentar simularla lo que hicieron fue mover los labios hacia abajo o contraer los músculos frontal y superciliar que son los que están ubicados en la parte anterior de las cejas.

La siguiente expresión en la lista fue el enojo con un porcentaje de reconocimiento de 36% correspondiente a 18 sujetos, bastante bajo para ser una expresión que en teoría es sencilla de simular, los sujetos lo que hacían era fruncir el ceño lo cual no era suficiente para una detección efectiva, para detectar el enojo normalmente se hacía con el ceño fruncido y la boca con una curvatura hacia abajo, Seguido estaba el disgusto con un porcentaje de éxito del 20% correspondiente a 10 sujetos, para la detección de esta expresión era necesario contraer los músculos de la nariz, las cejas y la curvatura de la boca hacia abajo, por lo que era de las expresiones más difíciles de lograr y se debe tener cuenta que no se les daba instrucciones a los sujetos para simular la expresión, sino que se dejaba a su criterio la simulación de la misma.

Finalmente, está la expresión de miedo con el porcentaje más bajo, pudiendo reconocer solo el 8% que corresponde a 4 sujetos y esto se debe a que la expresión se logra con la apertura de los músculos alrededor de los ojos, y la curvatura de las cejas y la boca hacia abajo como se observa en la figura 18.

Es importante resaltar que el diseño de estas pruebas no fue el más adecuado y que la forma correcta de realizarlas era tratando de generar el estímulo que provoca la expresión facial por medio de videos y actividades, adicionalmente **la capacidad de la visión artificial llega hasta donde somos capaces de diferenciar los humanos** y está es otra razón de los resultados tan variados y algunos muy poco precisos en el reconocimiento de expresiones faciales.

## 6.6 Edad y género

Es importante resaltar que durante el registro se debía diligenciar el género y la edad, pero los datos que aparecen en el recuadro después de la detección del rostro corresponden a las estimaciones realizadas por el algoritmo y no a los datos que se registraron. Esto con el fin de comparar las estimaciones con los valores reales del registro.

Respecto al reconocimiento de género se pueden observar los resultados en la tabla 4 donde se puede ver que de 50 sujetos tuvo una estimación acertada sobre 47 sujetos correspondientes al 94% de la muestra, hay que resaltar que los tres casos donde no hizo bien la estimación correspondían a mujeres que al momento del registro tenían el cabello completamente recogido, por lo que no quedaba en los puntos de detección del contorno del rostro y pudo ser uno de los factores que pudo haber afectado el resultado. Sin embargo, de forma general se puede decir que el parámetro es muy estable durante la grabación en tiempo real y hace estimaciones acertadas.

Finalmente se analiza el parámetro edad por medio del histograma mostrado en la figura 22 que relaciona a cada sujeto con el porcentaje de error de la edad estimada con respecto a la edad real, se puede observar que la mayoría de barras son rojas y componen los valores más altos del histograma, lo que quiere decir que el sistema tiende a poner una edad por debajo de la real con grandes porcentajes de diferencia y en menor medida estima una edad por encima de la real, pero cuando lo hace es con porcentajes mucho más bajos.

Es importante tener en cuenta que este es un parámetro difícil de estimar con precisión y que nosotros mismos como humanos tendemos a fallar tratando de hacer la misma estimación. Una forma de mejorar la prueba es **hacer el estudio con la percepción que tienen las personas sobre la edad de un sujeto** y comparar con las estimaciones resultados del sistema.

## 7. Conclusiones

El objetivo principal del proyecto era el desarrollo de un sistema de reconocimiento facial que fuese capaz de identificar expresiones faciales y estuviese asociado a una base de datos que relacionara la información personal de sujetos registrados. Se desarrolló un sistema funcional capaz de detectar un rostro en una imagen, extraer puntos de referencia y a partir de las características realizar una estimación sobre la edad, el género y la expresión facial, el sistema fue asociado a una base de datos creada en MySQL donde se almacenaban los datos personales y los datos biométricos de los rostros que posteriormente se usaban para realizar el proceso de comparación.

Se diseñó una interfaz gráfica de usuario que permitía usar una cámara para el ingreso de una imagen y retornar los datos asociados a un sujeto previamente registrado en la base de datos, adicionalmente, por medio de un video se observaba en tiempo real el nombre, la edad, el género y la expresión facial del rostro identificado, las pruebas de desempeño mostraron una excelente percepción de funcionalidad del sistema. Sin embargo, queda pendiente mejorar aspectos estéticos de la interfaz gráfica y modificar la ubicación de los botones para mejorar la experiencia de usuario.

Respecto a las pruebas de validación se encontró que para las expresiones faciales el sistema tiene una precisión del 98% identificando neutralidad y felicidad, para las demás expresiones faciales no es muy confiable y se debe tener en cuenta la naturaleza forzada de las pruebas. Identificando el género tiene una precisión del 94%, se debe tener en cuenta que la muestra de estudio fue relativamente pequeña y se deben ampliar el número de registros. Finalmente, el parámetro de estimación de la edad no es confiable ya que los porcentajes de diferencia con respecto a la real son muy altos y variables.

## Bibliografía

- [1] Ministerio de sanidad y política social (2010), “Unidad de urgencias hospitalarias, estándares y recomendaciones” Recuperado el 20 de julio de 2019 de: <https://www.mscbs.gob.es/organizacion/sns/planCalidadSNS/docs/UUH.pdf>
- [2] Murillo-Escobar, M. A., Cruz-Hernández, C., Abundiz-Pérez, F., & López-Gutiérrez, R. M. (2014). Cifrado caótico de plantilla de huella dactilar en sistemas biométricos. In Congreso Latinoamericano de Control Automático (pp. 18-23).
- [3] S. Domínguez, “Reconocimiento facial mediante el Análisis de Componentes Principales (PCA)” Tesis de grado, TSC, Universidad de Sevilla, Sevilla, 2017.
- [4] Wen, Y., Zhang, K., Li, Z., & Qiao, Y. (2016, October). A discriminative feature learning approach for deep face recognition. In European conference on computer vision (pp. 499-515). Springer, Cham.
- [5] TensorFlow (2019), An end-to-end open source machine learning platform, Recuperado del 20 de Julio de 2019 de: <https://www.tensorflow.org/>
- [1] Real Academia Española. (s. f.). En Diccionario de la lengua española (versión electrónica 23.3.). Recuperado de <https://dle.rae.es/inteligencia>
- [2] Rusell, S. y Norving, P. (1994). Artificial Intelligence A Modern Approach. Harlow, United Kingdom, Pearson.
- [3] Andreas Kaplan; ¿Michael Haenlein (2019) Siri, Siri in my Hand, who’s the Fairest in the Land? On the Interpretations, Illustrations and Implications of Artificial Intelligence, Business Horizons, 62(1), 15-25
- [4] E. Maduell, “Visió artificial,” Univ. Oberta Catalunya, vol. 1, 2012.
- [5] S. Domínguez Pavón, “Reconocimiento facial mediante el Análisis de Componentes Principales (PCA),” p. 74, 2017.
- [6] Bonsor, K., & Johnson, R. (2001). How facial recognition systems work. HowStuffWorks. Com Np, 4.
- [7] W. W., Bledsoe, “The Model Method in Facial Recognition,” Technical Report, PRI 15, Panoramic Research, Inc., Palo Alto, California, 1964.
- [8] Harmon, L. D., Hunt, W. F.: Automatic recognition of human face profiles. Computer Graphics and Image Processing, 6(2), pp. 135–156 (1977)
- [9] Kanade, T.: Computer recognition of human faces, Birkhäuser Basel, Vol. 47 (1977)
- [10] L. Sirovich and M. Kirby, Low-dimensional procedure for the characterization of human, Journal of The Optical Society of America A-optics Image Science and Vision, 4, 1987.

- [11] P. Viola and M. Jones, Rapid object detection using a boosted cascade of simple features, In Computer Vision and Pattern Recognition. Proceedings of the 2001 IEEE Computer Society Conference, 2001.
- [12] Zhao, W., Chellappa, R., Phillips, P. J., Rosenfeld, A.: Face recognition: A literature survey. ACM computing surveys (CSUR), 35(4), pp. 399–458 (2003)
- [13] Taigman, Y., Yang, M., Ranzato, M., & Wolf, L. (2014). DeepFace: Closing the Gap to Human-Level Performance in Face Verification. 2014 IEEE Conference on Computer Vision and Pattern Recognition. doi:10.1109/cvpr.2014.220
- [14] Greer, W. (2019). Identity Authentication Through Face Recognition, Los Angeles, CA, EE. UU.: <https://www.facefirst.com/blog/breve-historia-del-reconocimiento-facial/>
- [15] Gurovich, Y., Hanani, Y., Bar, O. et al. Identifying facial phenotypes of genetic disorders using deep learning. Nat Med 25, 60–64 (2019) doi:10.1038/s41591-018-0279-0
- [16] Díaz, G. (2012). FacePhi Beyond Biometrics. España: <https://www.facephi.com/es/noticias/sala-prensa/facephi-reinforces-its-worldwide-leadership-by-creating-a-zero-risk-security-product-addressed-to-the-banking-sector/>
- [17] IBM Knowledge Center. Nodos de modelado – Redes neuronales.: [https://www.ibm.com/support/knowledgecenter/es/SS3RA7\\_sub/modeler\\_mainhelp\\_client\\_ddita/components/neuralnet/neuralnet\\_model.html](https://www.ibm.com/support/knowledgecenter/es/SS3RA7_sub/modeler_mainhelp_client_ddita/components/neuralnet/neuralnet_model.html)
- [18] P. Larrangaña Y A. Moujahid. Redes neuronales. 1ra. Ed. Departamento de Ciencias de la Computación e Inteligencia Artificial Universidad del País Vasco–Euskal Herriko Unibertsitatea
- [19] Informática Aplicada a la Ingeniería de Procesos – Orientación I, Universidad Tecnológica Nacional – Facultad Regional Rosario Departamento de Ingeniería Química Grupo de Investigación Aplicada a la Ingeniería Química (GIAIQ) Rosario – Marzo del 2001
- [20] Loncomilla, P. (2016). Deep learning: Redes convolucionales.
- [21] K. Sung, C. Riley. Applications of convolution in Image processing with MATLAB. University of Washington, 2013.
- [22] Pacheco, M. A. (2017). Identificación de sistemas no lineales con redes neuronales convolucionales.
- [23] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich. Going Deeper with convolutions. in: CVPR, 2015.
- [24] Víctor Olaya (2018), Sistemas de Información Geográfica, Recuperado el 25 de diciembre de 2019 de: [https://volaya.github.io/libro-sig/chapters/Bases\\_datos.html](https://volaya.github.io/libro-sig/chapters/Bases_datos.html)
- [25] B. Gustavo. (2019, mayo 13). ¿Qué es MySQL? Explicación detallada (1ra ed.) [Online]. Disponible: <https://www.hostinger.co/tutoriales/que-es-mysql/>

- [26] Alimadadi, S., Mesbah, A., & Pattabiraman, K. (2016, May). Understanding asynchronous interactions in full-stack JavaScript. In 2016 IEEE/ACM 38th International Conference on Software Engineering (ICSE) (pp. 1169-1180). IEEE.
- [27] Castillo, A. A. (2017). Curso de Programación Web: JavaScript, Ajax y jQuery. IT Campus Academy.
- [28] TensorFlow. (2019), An end-to-end open source machine learning platform, Recuperado del 25 de Diciembre de 2019 de: <https://www.tensorflow.org/>
- [29] V. Mühler. (2018 Jun 25). face-api.js — JavaScript API for Face Recognition in the Browser with tensorflow.js Available: <https://itnext.io/face-api-js-javascript-api-for-face-recognition-in-the-browser-with-tensorflow-js-bcc2a6c4cf07>
- [30] P. Ruiz. (2018, Oct 8). Understanding and visualizing ResNets. [Online] Available: <https://towardsdatascience.com/understanding-and-visualizing-resnets-442284831be8>
- [31] Kazemi, V., & Sullivan, J. (2014). One millisecond face alignment with an ensemble of regression trees. 2014 IEEE Conference on Computer Vision and Pattern Recognition. doi:10.1109/cvpr.2014.241
- [32] Sagonas, C., Tzimiropoulos, G., Zafeiriou, S., & Pantic, M. (2013). 300 Faces in-the-Wild Challenge: The First Facial Landmark Localization Challenge. 2013 IEEE International Conference on Computer Vision Workshops. doi:10.1109/iccvw.2013.59
- [33] B. Blagojevic. (2019, Sep 2). How Facial Recognition Works Part 2, Facial Landmarks. [Online] Available: <https://medium.com/ml-everything/how-facial-recognition-works-part-2-facial-landmarks-72f1b0e2a33a>
- [34] V. Mühler. (2018 Jun 25). face-api.js — JavaScript API for Face Recognition in the Browser with tensorflow.js Available: <https://github.com/justadudewhohacks/face-api.js/tree/master/src/faceLandmarkNet>
- [35] V. Mühler. (2018 Jun 25). face-api.js — JavaScript API for Face Recognition in the Browser with tensorflow.js Available: <https://github.com/justadudewhohacks/face-api.js/tree/master/src/ageGenderNet>
- [36] V. Mühler. (2018 Jun 25). face-api.js — JavaScript API for Face Recognition in the Browser with tensorflow.js Available: <https://github.com/justadudewhohacks/face-api.js/tree/master/src/faceExpressionNet>