

Sistema de recomendación de marcos de gafas basado en inteligencia artificial: Eyewear consultant



Joseph Maverick Idrobo Cerón

Yehiner Ancizar Molina Zemanate

Dirigido por: Julián Fernando Muñoz Ordóñez

Corporación Universitaria Comfacaucá - Unicomfacaucá

Facultad de ingenierías

Ingeniería de sistemas

Popayán, Colombia

2022

DATOS GENERALES	
Facultad:	Ingeniería
Programa académico:	Ingeniería de Sistemas
Título del proyecto:	Sistema recomendador de marcos de gafas basado en inteligencia artificial: EyeWear Consultant
Grupo de investigación:	MIND – Grupo de Investigación en Computación e Informática Aplicada
Línea de investigación:	Semillero de Inteligencia Artificial & Big Data (IA&BD)
Proyecto institucional asociado:	
Fecha de aprobación del perfil:	
Fecha de elaboración de la propuesta:	21/08/2021
Fecha de radicación:	

DATOS ESTUDIANTE	
Nombres y apellidos	Joseph Maverick Idrobo Cerón
Cédula de ciudadanía	1061763342
Correo electrónico	josephidrobo@unicomfacauca.edu.co
Número de celular	3116206728
DATOS ESTUDIANTE	
Nombres y apellidos	Yehiner Ancizar Molina Zemanate
Cédula de ciudadanía	1061735918
Correo electrónico	yehinermolina@unicomfacauca.edu.co
Número de celular	3218483782

Nota de aceptación:

El director de la opción de grado y el jurado evaluador del trabajo han aprobado el presente documento de grado, cumpliendo con los requisitos exigidos por la Corporación Universitaria Comfacauca de Popayán, para optar al título de ingeniería sistemas.

Director opción de grado

Jurado evaluador

Contenido

CAPÍTULO I. CONTEXTO.....	9
1.1 Planteamiento del problema.....	9
1.2 Objetivos.....	12
1.2.1 Objetivo general.....	12
1.2.2 Objetivos específicos.....	12
1.3 Justificación.....	13
CAPÍTULO II: MARCO DE REFERENCIA.....	15
2.1 Marco teórico.....	16
2.1.1 Gafas.....	16
2.1.2 Historia de las gafas.....	17
2.1.3 Conjunto de datos.....	18
2.1.4 Interfaz de usuario.....	21
2.1.5 Red neuronal convolucional.....	24
2.1.5.1 Construcción de la red neuronal convolucional.....	25
2.1.6 Visagismo.....	26
2.1.7 Métricas de clasificación.....	31
2.2 Marco conceptual.....	32
2.2.1 Redes neuronales convolucionales para la clasificación de imágenes.....	32
2.2.2 Arquitecturas para la clasificación de imágenes.....	33
2.2.3 Arquitecturas para la clasificación de imágenes Alexnet.....	34
2.2.4 Arquitecturas para la clasificación de imágenes VGGNet.....	35
2.2.5 Arquitecturas para la clasificación de imágenes GoogLeNet.....	36
2.2.6 Arquitecturas para la clasificación de imágenes ResNet.....	37

2.2.7 Arquitecturas para la clasificación de imágenes ResNeXt	38
2.2.8 Arquitecturas para la clasificación de imágenes Trimps-Soushen	38
2.2.9 Arquitecturas para la clasificación de imágenes InceptionV3	39
2.3 Estado del arte	40
CAPITULO III METODOLOGÍA	46
3.1 Entrenamiento de la red neuronal	49
3.2 Código entrenamiento de RedNeuronal	51
CAPÍTULO IV RESULTADOS Y ANÁLISIS	57
4.1 Interfaz grafica	57
4.1.1 Evidencia del código y explicación.	59
4.2. Conclusiones	65
4.3 Trabajos futuros.	66
Referencias	67

Lista de tablas

TABLA I Actividades y resultados relacionados con el desarrollo del proyecto.	48
TABLA II Métricas de evaluación obtenidas durante el proceso de entrenamiento del modelo de red neuronal convolucional (2801 imágenes de entrenamiento)..	64
TABLA III Métricas de evaluación obtenidas durante el proceso de validación del modelo de red neuronal convolucional (1199 imágenes de prueba)..	64

Lista de figuras

Fig. 1 Ejemplo de las métricas (líneas punteadas color blanco) de visagismo sobre un Rostro tipo redond [17].	27
Fig. 2 Ejemplo de las métricas (líneas punteadas color blanco) de visagismo sobre un Rostro tipo ovalado [17].	28
Fig. 3 Ejemplo de las métricas (líneas punteadas color blanco) de visagismo sobre un Rostro tipo ovalado [17].	29
Fig. 4 Ejemplo de las métricas (líneas punteadas color blanco) de visagismo sobre un Rostro tipo cuadrado [17].	30
Fig. 5 Ejemplo de las métricas (líneas punteadas color blanco) de visagismo sobre un Rostro tipo rectangular [17].	30
Fig. 6 Matriz de confusión [22].	31
Fig. 7 Arquitectura de red neuronal convolucional para clasificación de imágenes [18].	33
Fig. 8 Arquitectura Alexnet [19].	35
Fig. 9 Arquitectura VGGNet [19].	36
Fig. 10 Arquitectura de red neuronal GoogLeNet [19].	36
Fig. 11 ResNet bloques residuales [19].	37
Fig. 12 Arquitecturas ResNeXt ampliación de Resnet [19].	38
Fig. 13 Arquitectura categoría con respecto a los errores [19].	39
Fig.14 ArquitecturaInception[20].	40
Fig. 15 Ejemplo de imágenes descargadas del dataset de Kaggle.	50
Fig. 16 Ejemplo de imagen recortada manualmente.	51
Fig. 17 Importan las librerías	51
Fig. 18 Se carga el data set.	52
Fig. 19 Parámetros iniciales del entrenamiento.	52
Fig. 20 Convertir en formato dataset.	53

Fig. 21 Se modifica el data set.	53
Fig. 22 Aumento de datos del data set.	54
Fig. 23 Se crea la arquitectura de la red neuronal convolucional.	54
Fig. 24 Configurar los checkpoints.	55
Fig. 25 Se compila el modelo.	55
Fig. 26 Entrenamiento	55
Fig. 27 Generación de métricas.	55
Fig. 28 Generación de métricas visuales	56
Fig. 29 Inicia el sistema en la captura de la imagen: Fuente propia.	57
Fig. 30 Rostro recortado: Fuente propia.	58
Fig. 31 recomendación de marcos: Fuente propia.	58
Fig. 32 Se crea el canvas y los botones de tomar fotografía y salir: Fuente propia	59
Fig. 33 Función capturar fotografía: Fuente propia.	59
Fig. 34 Define marcos a mostrar: Fuente propia.	60
Fig. 35 Muestra los marcos recomendados: Fuente propia	61
Fig. 36 Muestra los frames: Fuente propia	62
Fig. 37 Validar cámara encendida: Fuente propia	62
Fig. 38 Destruir ventanal salir: Fuente propia	63
Fig. 39 Capturar hora y fecha: Fuente propia.	63

CAPÍTULO I. CONTEXTO

1.1 Planteamiento del problema

La organización mundial de la salud (OMS) por lo menos 2200 millones de personas tienen ceguera o insuficiencia visual a nivel mundial, se estima que alrededor del 50% de dicho grupo, conociendo sus enfermedades visuales (miopía, hipermetropía, glaucoma, cataratas, entre otras), aún no han iniciado un tratamiento correctivo, lo que conlleva a que dichas personas realicen sus actividades rutinarias sin atender a su problema visual. La OMS manifestó que tanto el envejecimiento de la población, los cambios en los estilos de vida, los problemas y obstáculos al recibir atención oftalmológica son más evidentes en los países cuyos ingresos son considerados bajos y medios, lo anterior constituye los principales factores que generan un rápido incremento en la deficiencia visual afectando la calidad de vida de los pacientes. Es de aclarar, que un paciente con problemas visuales sin un tratamiento correctivo a tiempo se verá obligado en un futuro a invertir una cantidad de dinero mayor, ya sea en lentes de contacto o gafas oftálmicas con una tecnología más avanzada con el fin de mejorar su agudeza visual [1].

Un estudio que se realizó a trabajadores en España por la Sociedad de Prevención de Fremap empresa encargada de los accidentes laborales y enfermedades profesionales, evidenció que esta problemática afecta a más del 50% de las personas que laboran ya que esta mitad de personas sufre alguna de las enfermedades visuales o tiene defectos refractivos, en esta investigación se analizaron 780 mil historias clínicas realizadas a trabajadores en el año 2013 y cerca del 52% de los empleados usaban algún método temporal para corregir su visión y poder desempeñar mejor sus actividades laborales. Los errores de refracción son un trastorno ocular común que provoca visión borrosa y que se manifiesta principalmente de cuatro formas: miopía

(pérdida de visión lejana), hipermetropía (pérdida de visión cercana), astigmatismo (distorsión de las imágenes) y presbicia (visión cercana defectuosa, que en muchos casos impide leer sin gafas). Como consecuencia de la falta de servicios de corrección de los errores de refracción se estima que alrededor de 119 millones de personas tienen discapacidad visual, a pesar de que existen soluciones eficaces para el tratamiento de estos problemas la falta de asequibilidad afecta a la mayoría de los pacientes [2].

La OMS preocupada por los altos índices de insuficiencia visual ha hecho un llamado a través del informe mundial sobre la visión, de integrar la atención oftálmica dentro de la cobertura sanitaria. Esta solicitud forma parte de la prevención y control para evitar mayores afectaciones, aplicando campañas equitativamente para incluir a la población con difícil acceso, zonas rurales y remotas, con bajos recursos económicos, que se han visto afectadas debido a que presentan mayores dificultades de acceso a los servicios. En este sentido, el aspecto familiar también se realizan campañas donde incentivan a los padres a hacer más acompañamiento a sus hijos que recientemente han sido formulados con gafas, acompañándolos en su transición de uso, sin presiones, brindado espacios y confianza para que sea el niño quien empiece a usar y adaptarse a los lentes [2].

En Colombia se realizan campañas como “*veo bien aprendo bien*”, en el cual se implementan estrategias entre el sector salud y educación para identificar en edades tempranas alteraciones visuales y mejorar el acceso oportuno a los servicios de salud visual, prevención de ambliopía, prevención de ceguera por glaucoma, tecnología en salud visual, entre otras. Se realiza un seguimiento a los casos encontrados y se socializan en la comunidad escolar y padres de familia con el fin de sensibilizar y educar sobre las problemáticas que se presentan, permitiendo realizar acciones oportunas de prevención [3].

Sharezah Housden, un experto en anteojos de Culter y Gross of Knightsbridge habla que en la elección de las gafas no solo es importante el diseño de los marcos, sino que también se detalla la conveniencia del cabello y la forma de la cara. Por lo tanto, esta se convierte en la principal razón para investigar cómo hacer que sea más fácil y satisfactorio la elección de marcos para gafas buscando la adaptación a la forma de la cara en los pacientes con problemas visuales. Más de la mitad de las personas que laboran en Colombia tienen enfermedades visuales y la mayor parte necesita un tratamiento o una cirugía ocular para mejorar su calidad de vida, tener un éxito laboral y sentirse más atractivos y cómodos con su apariencia estética [4].

Algunas investigaciones han enfocado sus esfuerzos en generar estrategias para incrementar el uso y adaptación de los lentes en las personas con problemas de visión. En el aspecto familiar también se realizan campañas donde incentivan a los padres a hacer más acompañamiento a sus hijos que recientemente han sido formulados con gafas, para que hagan el apoyo emocional, pero evitando exceso de protección, permitiéndoles escoger sus gafas, acompañándolos en su transición de uso, sin presiones, brindando espacios y confianza para que sea el quien empiece a usarlos y continúe siendo un niño alegre y seguro de sí mismo [5].

La Fundación ALAIN AFFLELOU quiere concienciar a los padres y profesores, como prescriptores, de la importancia de revisar la vista a los niños al menos una vez al año para asegurarse que ningún problema visual afecte a los niños en su rendimiento escolar. Cada año, entre septiembre y octubre, la Fundación ALAIN AFFLELOU pone en marcha la Campaña Escolar en favor de la Salud Visual Infantil. El objetivo de la campaña es detectar de manera precoz los posibles problemas visuales en los niños de entre 5 y 7 años, edad a la cual pueden ser fácilmente corregidos. A los que lo necesitan, la Fundación ALAIN AFFLELOU les entrega gratuitamente unas gafas graduadas adaptadas a su vista (gafa completa: montura especial de campaña + cristales orgánicos monofocales blancos, sin tratamiento) [6].

Esta investigación pretende la creación de un sistema recomendador de marcos para lentes que pueda ser utilizado como una estrategia para apoyar a las empresas e instituciones encargadas de proporcionar los lentes oftálmicos mediante el sistema EYEWEAR CONSULTANT para mejorar la adopción de marcos en los pacientes cuando realicen la adquisición de estos [7]. Se identifica la necesidad de aportarle acompañamiento al usuario para que tome la decisión de adquirir unos marcos para lentes, así como también se evidencia que los clientes que reciben una consulta oftalmológica no siempre adquieren o compran los marcos en la óptica que les brindó el servicio, a pesar de la excelente asesoría y buen trato que se le brinda. Toda esta información que se genera a partir de la observación del comportamiento de los clientes y usuarios en la óptica ha sido fundamental para identificar la necesidad de desarrollar un sistema tecnológico moderno que genere recomendaciones a las personas de acuerdo a su tipo de rostro. Para tales fines, el presente trabajo plantea la siguiente pregunta de investigación: ¿Cómo construir un modelo de clasificación basado en inteligencia artificial que permita crear recomendaciones sobre estilos de marcos para lentes basado en características faciales?

1.2 Objetivos

1.2.1 Objetivo general

Construir un sistema para la recomendación de marcos para lentes basado en el tipo de rostro de las personas utilizando visión por computadora y redes neuronales convolucionales.

1.2.2 Objetivos específicos

- Definir un conjunto y tipo de datos establecido en el estado del arte del proyecto relacionado con la clasificación de formas de rostros de personas según su morfología y técnicas de visagismo para ser utilizado en algoritmos de clasificación.

- Construir un modelo de aprendizaje profundo basado en una red neuronal convolucional para la clasificación de tipos de rostros que permita generar recomendaciones hacia los usuarios relacionado con marcos para lentes.
- Evaluar el modelo de clasificación de tipos de rostros utilizando métricas de validación como la matriz de confusión, coeficiente Kappa y porcentajes de error.

1.3 Justificación

Los estudios de la OMS indican que hay un porcentaje elevado de personas que tienen problemas visuales. Aunque se están realizando campañas para reducir estas estadísticas a través de organizaciones gubernamentales, solo se están realizando enfocadas en la población infantil en la cual se prioriza la prevención sanitaria. Si bien la población infantil es importante, ya que una buena educación conlleva a adultos responsables, se realizan menos campañas para las personas adultas y no se les brinda un acompañamiento oportuno cuando presentan problemas visuales, no son informados de los beneficios que pueden obtener al usar las gafas a tiempo y los inconvenientes a futuro si no las usan. De la misma forma se evidencia que son pocas las personas que investigan por cuenta propia sobre los beneficios de usar gafas y los perjuicios de no usarlas [2].

Se ha identificado que la miopía, el astigmatismo, entre otros problemas visuales podría corregirse usando gafas o lentes a tiempo en niños, muchos de los casos se agravan con el paso del tiempo debido a que no usan adecuadamente las gafas o interrumpen su uso debido a aspectos estéticos [8]. El uso de gafas o lentes formulados luego de evidenciar algún problema visual frena en algunos casos el aumento de dioptrías y puede ayudar a recuperar la visión, en pacientes con

hipermetropía menor a 4 dioptrías se ha logrado que en una edad adulta ya no tengan que utilizar sus gafas mejorando la calidad de vida y justificando el uso de gafas a temprana edad evidenciando correcciones visuales [32].

La finalidad de este proyecto es generar recomendaciones basadas en estudios científicos e investigativos para apoyar en la toma de decisión en el momento de la adquisición de un marco para gafas por parte de usuarios, además el proyecto busca fortalecer mediante la construcción de una solución tecnológica la temática relacionada con la adopción de marcos por parte del paciente. El proyecto buscará desarrollar un sistema que permite identificar el tipo de rostro en tiempo real a través de redes neuronales, creando una comunicación inmediata y directa con el usuario sugiriéndole el tipo de gafas que mejor se adapte a su fisionomía logrando satisfacción en el usuario que lo motivará a continuar usándolas.

La enfermedad visual hoy en día se ha convertido en algo muy común, se puede evidenciar que los pacientes no están usando sus lentes oftálmicos por diferentes factores, ya sea porque no encuentran una buena asesoría al momento de hacer sus compras o la falta de conciencia respecto a la enfermedad visual que padecen, la finalidad del proyecto RECOMENDADOR DE GAFAS es que la población identificada con problemas visuales sean asesorados tecnológicamente buscando un gran beneficio ya que se busca con esta asesoría un grado de satisfacción alto al momento de finalizar la compra de lentes oftálmicos y queden a gusto con la manera en cómo se ve y se siente este objeto que para muchas personas se ha vuelto tan indispensable en la rutina cotidiana, también se podrá utilizar en brigadas estudiantiles en las cuales se buscará una mayor aceptación de las gafas por parte de los pacientes y estos se encuentren a gusto con la escogencia de sus marcos disminuyendo el bullying por no tener las gafas que se vean estéticamente bien, buscando que la enfermedad visual sea mayormente controlada y previniendo que está en pacientes ya diagnosticados aumente de una manera rápida y sin control.

Solo los seres humanos establecemos la belleza, lo hacemos desde nuestro razonamiento lógico, influido por el entorno cultural y social en el que nos encontremos, por consiguiente que la belleza define la estética como el conjunto de cualidades físicas de una persona, el cual contribuye al bienestar y equilibrio de la personalidad, Es por ello que la estética es uno de los factores importantes para la selección de marcos para gafas, se evidencia que los usuarios dedican tiempo en la búsqueda de unos marcos que cumplan sus estándares de belleza y se amolden con su personalidad, convirtiendo al usuario con mayor seguridad y confianza al sentirse a gusto con este accesorio que los acompaña en su día a día. [36]

El presente trabajo está organizado por capítulos en donde se encuentra como primer capítulo el contexto de la investigación, abordando así la problemática, objetivos, y justificación. En el segundo capítulo está compuesto del marco de referencia y el marco conceptual, antecedentes y marco legal, posteriormente se encuentra el capítulo de la metodología y para finalizar están los resultados y análisis de la investigación.

CAPÍTULO II: MARCO DE REFERENCIA

En este capítulo, se trata el marco teórico del trabajo, se describen a detalle los fundamentos teóricos en los que se basa el presente proyecto de investigación, así como, sus características generales, objeto de estudio y la normatividad vigente que se relaciona con la actividad científica a desarrollar.

2.1 Marco teórico

Los siguientes son los principales conceptos que se usaran en el desarrollo del presente proyecto.

2.1.1 Gafas

Es una noción que alude a los anteojos cuya sujeción se realiza por detrás de la cabeza o aprovechando la estructura de las orejas. Los anteojos, en tanto, son instrumentos ópticos que incluyen dos lentes para favorecer la visión de la persona que los utiliza. Es por ello por lo que en la mayoría de los casos consideraron que el concepto se emplea en plural: gafas[30]. Es fundamental considerar que, en el lenguaje coloquial, gafas, anteojos y lentes son conceptos que se usan como sinónimos, aun cuando se destacan en diferentes zonas de habla hispana. De allí que los anteojos constan de un armazón o montura que sujeta a los anteojos. El armazón dispone de 2 patillas que se secundan en las orejas y de un arco que se localiza sobre la nariz. Tal cual, las gafas quedan justo delante de los ojos del cliente, que de esta forma puede arreglar deficiencias de su perspectiva, como el astigmatismo o la miopía.

Asimismo, los armazones de los lentes originalmente eran realizados en acero o hierro. Con la época, los productores comenzaron a usar resinas. Actualmente, es usual que se recurra a aleaciones de titanio o aluminio. Referente a los anteojos, tenemos la posibilidad de encontrarlos fabricados con cristal, polímeros plásticos o policarbonato. Más allá de los anteojos cuyos lentes permiten la corrección de deficiencias de la perspectiva, además hay otros tipos de gafas. Los lentes de sol muestran lentes oscuros para bloquear los relámpagos ultravioletas y defender a los ojos de la luz solar. Además, hay gafas de defensa que se aplican en ciertos oficios e inclusive en el campo del deporte[30].

Los lentes de sol “son uno de los equipos más frecuentes en épocas de calor, aun cuando en ciertas piezas de todo el mundo se aplican durante todo el año” [10]. Se habla de un producto que resulta bastante eficaz en verano para impedir que la luz del sol entorpezca el movimiento por la vía pública, empero además sirve para defender los propios ojos una vez que se está al aire independiente. A lo extenso de las vacaciones de verano en la playa, ejemplificando, la mayor parte de la población los usa.

2.1.2 Historia de las gafas

El invento de los lentes se atribuye a 2 monjes franciscanos, Alejandro de la Espina y Roger Bacon, quienes hacia medio siglo XIII comunicaron el secreto de la construcción de las lentes. Debido a su hallazgo, los frailes de más edad podían avanzar con sus tareas aun cuando tuviesen inconvenientes de perspectiva gracias a la vejez. El primer cuadro de una persona con gafas data del año 1.352, un retrato del Cardenal Hugo de Provenza leyendo en un escritorio, que ha sido pintado por Tomás de Módena. Esta fotografía está en el templo de San Nicolás de Treviso en Italia [3] . De allí se tiene la posibilidad de mencionar que Italia, precisamente Murano, en el siglo XIII, ha sido la cuna de la construcción de lentes. Las fábricas de cristal de Murano eran las únicas capaces de crear vidrio moldeable. Las primeras gafas fabricadas ahí fueron para la presbicia con una lente esmerilada convexa y un siglo después llegaron las lentes cóncavas para la miopía.

Después de cierto tiempo los lentes no padecieron cambios hasta 1.452, una vez que Gutemberg había inventado la imprenta y se produjo una revolución en la lectura de libros. La demanda de gafas incrementó y ha sido en aquel instante una vez que la construcción de lentes abandonó de ser un arte monacal para comenzar a producirse de manera masiva en los primeros talleres. Es de mencionar que otro gran hito en la narración de los anteojos llega durante el siglo XVIII, una vez que el de Estados Unidos Benjamin Franklin, más habitual por sus estudios sobre fenómenos eléctricos, fabricó las primeras lentes bifocales que permitían ver de cerca y de lejos.

Finalmente, a inicios del siglo XX los anteojos adoptaron la manera que actualmente conocemos, dejando atrás los modelos rudimentarios, elaborados en madera, cuero o huesos. Actualmente disponemos de un sinfín de diseños, maneras y colores a nuestro alcance [30].

2.1.3 Conjunto de datos.

“Es un conjunto de datos o archivos que alberga información que se procesa por medio de aplicaciones expertas en el procesamiento de datos ya que estos contienen gran cantidad de estadísticas e información, encontramos 4 tipos de conjunto de datos [11]:

- a) Archivo: estos se caracterizan por ser seguros, rápidos y eficientes.
- b) Folder: aquí se almacena una gran cantidad de información en diferentes conjunto de datos, estos tienen que estar en el mismo formato para que sean compatibles entre sí y se pueda hacer un análisis extenso de los datos.
- c) Base de datos: estos se utilizan en programas específicos de acuerdo con el formato en el que estén creados los datos almacenados.
- d) Web: son datos que están almacenados en una página web y se representan en URL donde almacena toda la información del sitio [11].

Un grupo de datos tiene los datos que el servidor ha cargado y procesado. Los pasos relacionados en la carga y el procesamiento de los datos por el servidor componen el proceso de creación del grupo de datos. A lo largo de la obra del grupo de datos, el servidor lee los datos de procedencia de los principios de registro, aplica transformaciones a campos de datos específicos y define las magnitudes ampliadas que se crearán desde los campos transformados. El proceso de creación se genera en 2 etapas: Procesamiento de registros y Transformación. Una vez construido el grupo de datos, puede usar las magnitudes ampliadas del grupo de datos para generar métricas y magnitudes derivadas para sus objetivos de estudio específicos [12]. La construcción de conjuntos de datos es considerada por González como:

“Un proceso donde se seleccionan los datos (las materias primas) que se usarán para generar el grupo de datos y se definen las transformaciones de datos (los pasos del proceso) que manipulan la información disponible en los datos para producir magnitudes ampliadas (los productos fabricados)” [12] .

2.1.3.1 Construcción de un conjunto de datos.

De lo anterior, se dice que la configuración del grupo de datos se refiere al proceso de versión de los archivos de configuración cuyos límites dan las normas para la obra del grupo de datos. El grupo de datos construido radica físicamente en el documento temp db guardado en los equipamientos servidor, sin embargo, los archivos de configuración del grupo de datos residen en un directorio para un perfil y por ello un perfil tiene un grupo de archivos de configuración que construyen un grupo de datos (incluidas sus magnitudes ampliadas) para un objetivo de estudio específico. Además, un perfil tiene las definiciones de entidades como métricas, magnitudes derivadas, espacios de trabajo, informes y observaciones que permiten a los analistas interactuar con el grupo de datos y obtener información de él.

De igual manera, el perfil cuyos archivos de configuración de conjuntos de datos está editando se llama perfil de conjuntos de datos. Un perfil de grupo de datos se refiere a diversos perfiles heredados, que tienen la posibilidad de ser cualquier perfil que piense y mantenga para que logre configurar la aplicación de Adobe para que se acomode mejor a sus necesidades de estudio. Un perfil de grupo de datos además puede hacer alusión a perfiles internos que se otorgan con la aplicación de Adobe para conformar la base de toda la funcionalidad disponible en la aplicación [12].

Asimismo, una transformación es un grupo de indicaciones que puede conceptualizar para sustraer o manipular información en los datos de acontecimiento. Cada transformación que defina se va a aplicar a cada registro de datos de acontecimiento (entrada de registro) para actualizar los campos de registro existentes o crear nuevos campos. Los resultados de las transformaciones se usan junto con las condiciones de ingreso de registro para evaluar qué entradas de registro se filtran fuera del grupo de datos a lo largo del procesamiento del registro. Es importante señalar que no todos los tipos de transformaciones tienen la posibilidad de usar a lo largo de la etapa de procesamiento de registros del proceso de creación del grupo de datos, por ello es necesario que se experimenten estos procesos al momento de realizar una construcción de datos:

Filtrado de registros: El grupo de datos tiene diversos límites usados para filtrar los datos que salen de las transformaciones. El filtrado se usa para especificar qué entradas de registro se usan en los pasos de procesamiento siguientes. Ejemplificando, los filtros tienen la posibilidad de definirse por, intervalo de tiempo, el estado de la contestación del servidor o la dirección IP y la información de usuario-agente. El Log Entry Condition es una prueba de filtrado personalizable. La prueba busca ciertas condiciones en los campos de cada ingreso de registro para decidir si la ingreso debería avanzar en el proceso de creación del grupo de datos. Si un acceso de registro no cumple la condición, el acceso se remueve del proceso de creación.

Identificación de campos para transformación: Si se va a pasar un campo de datos de la etapa de procesamiento de registros a la etapa de transformación para un procesamiento siguiente, debería identificarlo a lo largo del procesamiento del registro. Este requisito se aplica independientemente de si el campo está disponible a partir de las fuentes de registro o se crea desde transformaciones de datos aplicadas a los datos a lo largo del procesamiento del registro. Finalmente se puede concluir diciendo que, a lo largo de la etapa de transformación de la obra del grupo de datos, el procesamiento se genera en los datos agrupados y ordenados que se crean desde el procesamiento del registro. Se hacen transformaciones de datos extras y se generan magnitudes

de datos ampliadas para su uso en los estudios. También a lo largo de la transformación, puede entrar una muestra estadística de los datos que se hace más enorme mientras la etapa de transformación se acerca a completarse.

Los conjuntos de datos son un conjunto de datos o archivos que alberga información que se procesa por medio de aplicaciones expertas en el procesamiento de datos ya que estos contienen gran cantidad de estadísticas e información, encontramos 4 tipos de conjunto de datos [24]:

2.1.4 Interfaz de usuario

La interfaz de usuario o UI (User Interface) es un criterio que encierra arquitectura de información, patrones y diferentes recursos visuales que permiten interactuar de manera eficaz con sistemas operativos y softwares de varios dispositivos. Según Insfrán se trata del “medio que posibilita interactuar y comunicar la computadora portátil, smartphone, tableta y varios otros dispositivos que son parte del día a día de habitante del siglo XXI” [13] . Fácilmente, se puede conceptualizar por medio del cual el usuario interactúa con un dispositivo tecnológico. Esto encierra todos los puntos de vista de contacto entre el individuo y los accesorios. Para entender mejor, se debería estar consciente de que cada vez que se abre el correo electrónico, se hace una transferencia electrónica, se envía un emoticono o cualquier otra acción parecida se está comunicando con un programa y sistema operativo [13]. No se puede dejar de lado que existen 3 tipos de interfaz de usuario según su diseño y objetivo:

- Interfaz de hardware: Encierra todos esos recursos que permiten ingresar, procesar y dar datos, como los célebres teclados y “ratones”, así como las pantallas.
- Interfaz de programa: Es aquella que ofrece datos sobre los procesos y herramientas de control, la cual podría ser vista de forma sencilla por el usuario en la pantalla de su dispositivo.

- Interfaz software-hardware: En la interfaz del usuario, esta es la que sirve de puente entre la máquina y los individuos.

De todo lo anterior se puede decir que todo esto produce un círculo virtuoso en el que el dispositivo comprende las indicaciones del sujeto y, paralelamente, este entiende la información del artefacto, por medio de la traducción del código binario a recursos básicos. Además, la interfaz del usuario puede clasificarse según la manera en que interactúa con el ser humano. Bajo este esquema, se puede asegurar que hay interfaces de línea de comandos, gráficas de usuario y natural de usuario. La interfaz de línea de comandos es alfanumérica, o sea, solo muestra escrito, en lo que las interfaces gráficas del usuario representan visualmente los recursos de control y medida para optimizar la comunicación con el usuario y, generalmente, la vivencia de este[13].

2.1.4.1 Diseño de la interfaz de usuario

El diseño de interfaz de usuario se refiere al diseño estético de la interfaz el cual prioriza la experiencia visual del usuario. Una buena interfaz debe reducir el esfuerzo que hace el usuario, es por ello que debe cumplir las primordiales propiedades: [34]

- Claridad: Una buena interfaz transmite la información de forma rigurosa para eludir que el usuario cometa errores a lo largo de la relación.
- Concisión: Hablamos de darle al usuario solo la información que requiere y exige.
- Coherencia: Esta característica es la que hace intuitiva una interfaz, permitiéndole al individuo generar patrones de uso de forma simple y práctica.
- Flexibilidad: Una buena interfaz además se caracteriza por permitir que el usuario restaure recursos y deshaga ocupaciones.

Llamativo visual: Una vez que hablamos de vivencia del usuario, jamás puede subestimarse el valor del llamativo visual. Además de tener enormes funciones, una buena interfaz es agradable a la vista. Por medio de estas propiedades se consigue que el usuario realmente tenga una conexión

con cualquier dispositivo tecnológico y, desde luego, lo prefiera frente a otras alternativas semejantes. De allí que el diseño de la interfaz del usuario debe estar desarrollado bajo elevados estándares de calidad. Por cierto, hay fronteras de todo el mundo y de todo el mundo sobre las buenas prácticas para desarrollar este factor [34].

Tal vez, el estándar más habitual y seguido es la regla ISO 9241, que se concentra en la ergonomía de la relación entre el individuo y el sistema, especialmente en puntos como la facilidad de la comunicación y el dinamismo. En una página web, aplicación o cualquier herramienta digital, el diseño de la interfaz del usuario debe asegurar tolerancia a los errores, personalización, control, adecuación al aprendizaje y, más que nada, conformidad con las expectativas del usuario. Además, el diseño de una interfaz busca la adecuación a la labor y la auto descripción. Asimismo, hay estándares de coherencia y calidad gráfica que tienen que seguirse para desarrollar una interfaz, como ocupar códigos de colores para ayudar las labores de los usuarios, minimalismo y estandarización. Por esto, referente a la vivencia del usuario se refiere a la forma en que una persona ve y siente su relación con cualquier canal en línea, primordialmente una página web o blog.

Entre otras cosas, este criterio es sustancial para las tácticas de Marketing Digital. Además de producir contenidos de calidad y estrategias de atracción, esta metodología debería estar acompañada de diseños y programaciones simpáticas, intuitivas y eficientes. Claramente, es decir lo cual relaciona la vivencia del usuario con la interfaz del usuario [13]. Si la interfaz no es de calidad y carece de las propiedades que se han explicado con anterioridad, desafortunadamente la vivencia no va a ser gratificante. Frente a otros términos es fundamental dejar claro que una buena interfaz da la vivencia de relación exitosa e, inclusive, natural. Las superiores interfaces son esas que realizan olvidar que se está interactuando con una máquina, dispositivo o aplicación. Sin embargo, presentan que se está comunicando con un componente que forma parte imprescindible de las vidas y que es tan real y eficaz como las manos y los sentidos.

2.1.5 Red neuronal convolucional.

En los últimos años, las redes neuronales convolucionales tuvieron una enorme optimización en su funcionamiento, tanto de esta forma que fueron varios los trabajos que se hacen y publican todos los años, por todo el planeta. Su uso ha dejado de ser exclusivamente académico y objeto de análisis y averiguación a estar plenamente implementada en el día a día. Un caso muestra de eso son las novedosas aplicaciones que emergen sobre cada una de esas que se aplican para reconocer objetos, cosas, personas, fauna o flora mediante una imagen. Aun cuando las organizaciones no han quedado atrás debido a que comenzaron a apostarle a la evolución de las redes neuronales lanzando al mercado productos nuevos tanto en programa como hardware. De ahí que las mismas sean vistas como un tipo de Red Neuronal Artificial con aprendizaje supervisado que procesa sus capas imitando el córtex visual del ojo humano para detectar diversas propiedades en el acceso que definitiva hacen que logren detectar objetos [14]. Las labores primordiales que hacen las CNN según Pérez son:

- Detección/categorización de objetos.
- Categorización de escenas.
- Categorización de imágenes generalmente.

Las capas convolucionales de acuerdo con lo expresado se comprenden como “un grupo de filtros habitualmente denominados: campos receptivos, éstos se ajustan para la sustracción de propiedades de una señal” [14]. A diferencia de una red neuronal clásico, en la que cada neurona de una capa se conecta a cada una de las neuronas de la otra, conocido como Fully Connected (FC), en las CNN se comparten las neuronas por medio de filtros que permiten sustraer información de las imágenes de ingreso. Igualmente, cada capa de la CNN es un bloque con 3

primordiales cambiantes: ingreso, pesos y salida, en estas capas se muestra una característica importante como donde la salida de una capa se convierte en el acceso de la siguiente. Este proceso es secuencial y podría ser no lineal, en cada capa se hace una funcionalidad específica. Las CNN tiene como primordial labor sustraer patrones, en esta situación sobre imágenes, gracias a ello usan un entrenamiento no supervisado.

2.1.5.1 Construcción de la red neuronal convolucional.

La información procesada por el ser humano es percibida por la corteza visual, para que la visión pueda identificar un elemento se debió memorizar, aprender o conocer con anterioridad ese elemento. De esta misma forma la Red Neuronal Convolucional CNN debe aprender antes de identificar un elemento.

Para aprender la CNN hace lectura de las imágenes agrupadas por elemento en un conjunto de datos o dataset, a mayor número de imágenes que integren el dataset, el porcentaje de aprendizaje será más alto, cada imagen le brindará mayor información característica propias del elemento y así poder generalizarlo.

La CNN recibe una imagen, inmediatamente la normaliza, es decir transforma los píxeles de la imagen que van de 0 a 255 y los convierte en valores de 0 y 1. Una vez normalizado se hacen las convoluciones, las convoluciones podemos interpretarlas como filtros que nos permiten extraer las características importantes de la imagen, para ello el filtro recorre la imagen de izquierda a derecha, de arriba abajo y genera una nueva capa o capas que son conocidas como las capas ocultas. Es de resaltar que entre más filtros se apliquen, mayores características de la imagen son obtenidas. Sin embargo, antes de cada filtro es importante reducir el tamaño de las imágenes filtradas guardando las características más importantes, para realizar esta tarea de “subsampling” se ejecuta una operación conocida como Max-Pooling.

La operación max-pooling recoge la capa de la misma forma que lo hace el filtro o kernel en un rango establecido que por lo general es de 2x2, es decir hace un recorrido de 2 píxeles de

alto por 2 píxeles de ancho, y de ese rango se extrae el valor máximo o más alto, al final del recorrido la imagen resultante será a la mitad pero con la información más importante.

De esta forma se puede aplicar una cantidad determinada de convoluciones y se obtienen más características que permitirán conocer formas más complejas.[35]

2.1.6 Visagismo

El visagismo es un término que es fruto de la suma de 2 recursos relevantes: el término francesa “visage”, que significa “rostro”, y el sustantivo latino “ismo”, que es equivalente a “doctrina”. De allí que, por consiguiente, logre definirse como la ideología o práctica del rostro. De acuerdo con Hallawell el concepto de visagismo se emplea dentro de la estética para “hacer alusión a una conjunción de técnicas que busca exaltar la belleza del rostro por medio del maquillaje y el peinado” [15]. De allí que el autor del visagismo haya sido Claude Juillard, un estilista francés que tomó la decisión de centrarse en el análisis de las líneas, el volumen y las maneras del rostro. Desde dichos datos, el visagismo ofrece cómo maquillar y peinar al individuo según sus propiedades físicas y sus anhelos.

En este mismo orden, el visagismo apela a un espejo tridimensional para que la dama logre mirar cómo el estilista va desarrollando su trabajo. De esta modalidad, consigue comprender hacia dónde avanza su cambio de look e inclusive puede sugerir modificaciones. Según Juillard, el espejo tridimensional de su invención aumenta la satisfacción de las clientas de un estilista debido a que ofrece la probabilidad de mirar los cambios mientras se concretan. Ante esto es fundamental resaltar que el visagismo parte de ciertos preceptos que permanecen asociados a las cualidades del rostro y por ello se debe tener en cuenta que al momento de poder lograr que el visagismo ofrezca el resultado estético deseado, se debe considerar otra serie de puntos relevantes. En específico, se debe conseguir la igualdad entre los aspectos faciales. De allí que se tengan en importancia tanto

las magnitudes de la nariz (longitud, forma, perfil) como la distancia entre los ojos o lo cual es la manera de las cejas [15].

El visagismo nace como método de comunicación y análisis estético para realzar la belleza de cada persona. Es el conjunto de técnicas que ayudan a realizar un estudio detallado de las formas, volúmenes y líneas del rostro para adaptar sus posibilidades de maquillaje y estilismo. Fue el artista y estilista francés Claude Juillard quién se dio cuenta de que la formación de peluquería tenía una gran falla: no se explicaban bien los efectos de un corte en el rostro, se hacía lo que el cliente pedía sin informarle o cuestionarle, ni ofrecerle opciones más adecuadas a su rostro. Esto explicaba, según él, por qué gran cantidad de clientes salían descontentos de la peluquería, ya que no a todos les sentaba igual el mismo corte o peinado de moda. En base a eso crea el visagismo, método que se propone estudiar la forma del rostro para darle valor a la belleza de una persona con la ayuda de técnicas del estilismo, maquillaje y accesorios. Una herramienta muy utilizada en los salones más famosos y vanguardistas de peluquería y estética en general [16].

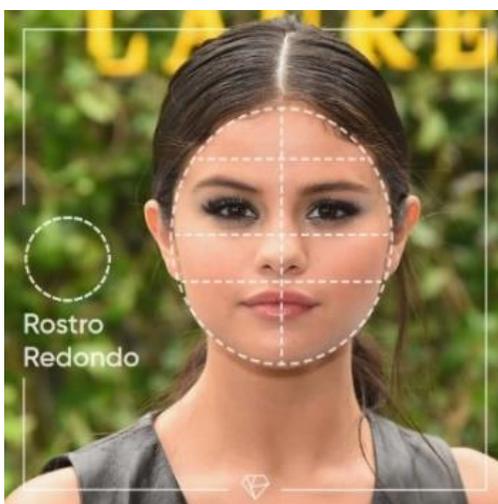


Fig. 1. Ejemplo de las métricas (líneas punteadas color blanco) de visagismo sobre un Rostro tipo redondo [17].

En la figura 1, se observa un tipo de rostro redondo que tiene la línea de los pómulos más ancha que la de las esquinas de la frente y la de los huesos de la quijada, en algunas personas es más ancha que en otras. Así mismo, la parte inferior del rostro tiende a formar un semicírculo y la línea que va de la frente a la barbilla es similar en distancia a la línea de los pómulos, es decir, el largo y ancho de la cara suelen ser muy similares en distancia. Los pómulos suelen ser la facción más prominente en las mujeres de cara redonda, y en general los ángulos laterales de la quijada que están justo debajo del oído no son muy prominentes.



Fig. 2. Ejemplo de las métricas (líneas punteadas color blanco) de visagismo sobre un Rostro tipo ovalado [17].

En la figura 2, Se observó que el tipo de rostro ovalado el cual es el más común de todos, tiene como característica que el ancho de las esquinas de la frente y el ancho de los huesos de la quijada es similar y si trazas una línea horizontal a la mitad del rostro, la distancia de la mitad del rostro a donde comienza la frente, así como la distancia de la mitad del rostro a donde termina la barbilla es muy parecida o igual. La distancia que hay entre los pómulos de las personas con cara ovalada es la medida más ancha de este tipo de rostro entre los ejes horizontales a comparación

del ancho de la frente y de la mandíbula que suelen ser ligeramente menos anchos lo cual hace que el rostro forme una silueta ovalada. Sin embargo, la línea más larga que hay en una cara ovalada es la línea vertical de frente a la barbilla.



Fig. 3. Ejemplo de las métricas (líneas punteadas color blanco) de visagismo sobre un Rostro tipo ovalado [17].

En la figura 3 se observa el rostro corazón que es junto con el rostro ovalado, uno de los más armoniosos. La característica principal de la cara en forma de corazón es el “pico de viuda”, que es una extensión en el crecimiento del cabello en forma de “V” justo en la parte central de la frente. Las personas con rostro de corazón suelen tener la frente más ancha que la línea de la quijada y el rostro termina con una barbilla puntiaguda dándole al rostro una semejanza a la silueta de un corazón.

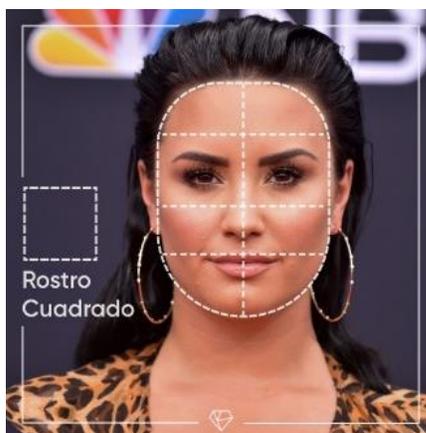


Fig. 4. Ejemplo de las métricas (líneas punteadas color blanco) de visagismo sobre un Rostro tipo cuadrado [17].

En la figura 4 se puede analizar que el rostro cuadrado es fácil de distinguir ya que el ancho de la frente, de los pómulos y de la mandíbula es casi igual, formando una línea vertical recta o casi recta entre estos tres puntos. Asimismo, esta distancia del ancho es similar a la distancia de la línea vertical que va desde la frente a la barbilla. Los ángulos de la mandíbula se distinguen por ser afilados y también es común que este tipo de rostro tenga pómulos muy prominentes.

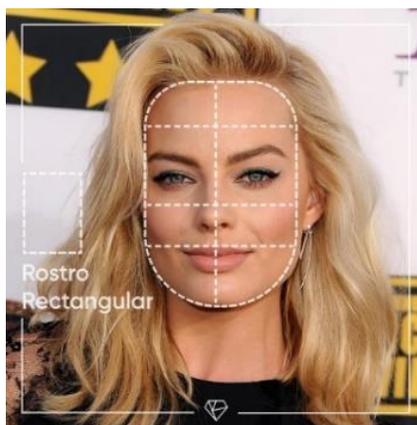


Fig. 5. Ejemplo de las métricas (líneas punteadas color blanco) de visagismo sobre un Rostro tipo rectangular [17].

En la figura 5 se puede observar que una característica no tan visible en una cara cuadrada son las curvas, este tipo de rostro es principalmente anguloso [17]. Finalmente se debe resaltar que el visagismo no sólo se recurre al maquillaje y a la peluquería sino además a la colorimetría. Bajo este último término se hace referencia a que para lograr conseguir la imagen que se quiere se debe tener en importancia la tonalidad de la dermis y además el color de ojos, al igual que se debe tener en cuenta que donde cobra particular protagonismo el visagismo es dentro del entorno del show, precisamente de la interpretación. Y es que los actores para lograr interpretar a diversos personajes, con una personalidad distinta, se someten a sesiones de maquillaje, peluquería y colorimetría. Debido a aquellas es como logran siendo la misma persona poder ofrecer vida a figuras que nada deben ver entre sí ni por edad, doctrina o maneras de comportamiento.

2.1.7 Métricas de clasificación.

Para la evaluación de este proyecto se utilizaron las siguientes métricas de clasificación: matriz de confusión y coeficiente Kappa.

Una matriz de confusión representada en una tabla con cuatro combinaciones distintas donde se muestran valores reales, los datos predichos que son aquellos que proporciona el modelo de red neuronal luego de haber realizado el respectivo aprendizaje con la información ingresada [21].

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Fig. 6 Matriz de confusión [22].

El coeficiente kappa es una medida estadística para elementos cualitativos. Se puede decir que es una medida mayor a la del cálculo del porcentaje de concordancia, ya que tiene en cuenta que la concordancia se produce por casualidad. Teniendo en cuenta que kappa es una medida más conservadora, su uso en la selección de características de envoltura es adecuado para probar el rendimiento de los modelos [23].

2.2 Marco conceptual

2.2.1 Redes neuronales convolucionales para la clasificación de imágenes

Las redes neuronales son muy útiles para, a partir de un dato de entrada, hacer una clasificación o regresión a partir de un modelo entrenado. Esta misma idea se puede aplicar a una imagen, para obtener, por ejemplo, qué representa la imagen (clasificación) o detectar un objeto que forme parte de ella (segmentación, detección, identificación, etc). Para ello se recurre a las redes neuronales convolucionales (CNN, por sus siglas en inglés), las cuales han supuesto una revolución en el sector del reconocimiento de imágenes, ya que, a diferencia de las redes neuronales convencionales y otros algoritmos de clasificación de imágenes, usan un procesamiento relativamente pequeño [18].

La red es entrenada a partir de imágenes utilizando solo píxeles y etiquetas como entradas. Aprenden usando filtros y aplicándolos a las imágenes. El algoritmo toma un cuadrado pequeño (o ‘ventana’) y comienza a aplicarlo sobre la imagen. Cada filtro permite que la CNN identifique ciertos **patrones** en la imagen. La CNN busca partes de la imagen donde un filtro coincida con el contenido de la imagen. En este caso las CNN toman como entrada tensores de forma, que incluyen alto y ancho de la imagen, en píxeles, y el número de canales de color (1, si

la imagen es en blanco y negro, o 3, si es en color RGB. Como salida se obtendría la clase o la probabilidad de que la imagen pertenezca a una clase particular. (Ver Figura. 7) [18].

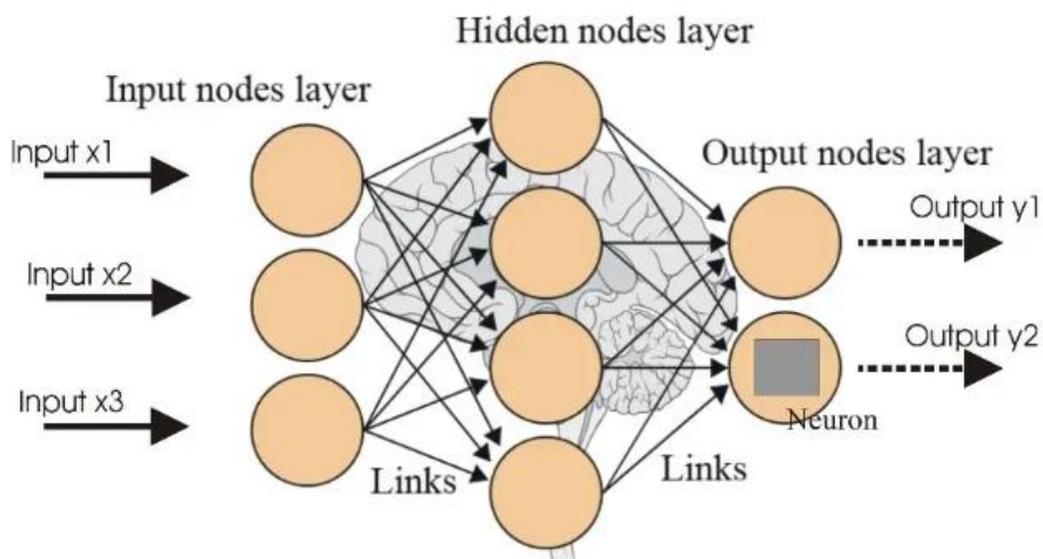


Fig. 7. Arquitectura de red neuronal convolucional para clasificación de imágenes [18].

2.2.2 Arquitecturas para la clasificación de imágenes

La clasificación de imágenes ha sido una de las principales áreas de investigación dentro del campo de visión artificial. Desde hace algo más de una década ha estado dominado por las redes neuronales convolucionales, CNN, que a pesar de los buenos resultados y grandes éxitos que han cosechado, sufren de varios inconvenientes. Para solventar estos problemas, recientemente se ha desarrollado un nuevo algoritmo de redes neuronales llamado redes de cápsulas (*CapsNet*) basado en vectores y en un nuevo algoritmo de entrenamiento llamado algoritmo de enrutamiento

dinámico, de manera que podría mejorar los resultados de conocidas arquitecturas para la clasificación de imágenes como lo son las redes neuronales convolucionales.

En este estudio se ha probado el rendimiento de esta nueva arquitectura, y se ha comparado con una arquitectura convolucional basada en bloques residuales. Para contrastar los resultados se han usado conjuntos de datos de distintas complejidades (MNIST, Fashion-MNIST, COIL-100 y CIFAR-10). A pesar de la buena perspectiva teórica que ofrece este nuevo modelo en la teoría, los resultados obtenidos todavía no son lo esperados en cuanto se manejan conjuntos de datos con cierta complejidad. Con MNIST y muy cerca, Fashion-MNIST sí se han mejorado los resultados, mientras que con los conjuntos de datos más complejos (COIL-100 Y CIFAR-10), los resultados no han sido los esperados [19].

2.2.3 Arquitecturas para la clasificación de imágenes Alexnet

En el año 2012, con el modelo *AlexNet*, fue la primera vez que ganó el reto una arquitectura basada en CNN y la tasa de error experimentó una bajada considerable. Consiguió que la tasa de error que estaba por encima del 25% pasara a ser menor del 17%. *AlexNet* tiene una arquitectura mucho más compleja y profunda que las redes que se habían usado hasta el momento, como *LeNet*. Estaba compuesta por 5 capas convolucionales y 3 capas totalmente conectadas (FC *fully connected*) como se observa en la figura 7, con 60 millones de parámetros y 650000 neuronas. El entrenamiento duró 6 días [19].

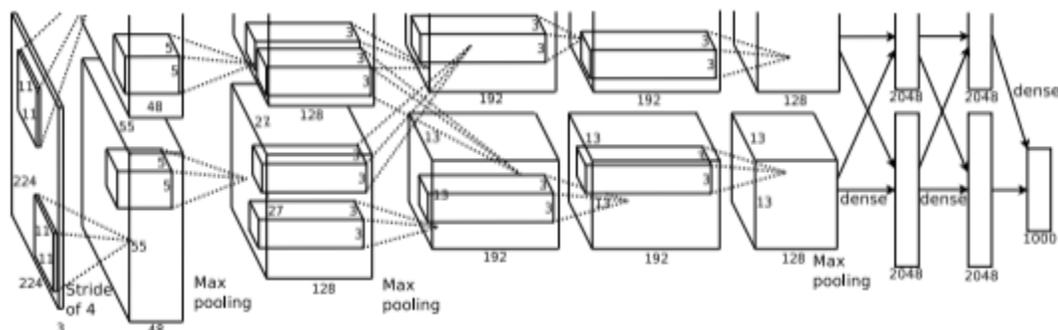


Fig. 8. Arquitectura Alexnet [19].

2.2.4 Arquitecturas para la clasificación de imágenes VGGNet

Vale la pena mencionar al finalista del reto en 2014, que fue *VGGNet*. Esta arquitectura ha sido una arquitectura muy influyente, ya que reforzó el razonamiento intuitivo detrás de las CNN de que para que una de estas redes logre una representación jerárquica de los datos suficientemente buena, la red convolucional tiene que tener una estructura de capas profunda. Además, esta red también destacó por su simplicidad: está formada por 19 capas convolucionales (detrás de cada cual se aplica la función *ReLU*) con filtros de tamaño 3x3, con stride y pad a 1, y las capas de *max pooling* (2x2) con *stride* a la configuración de los pesos está disponible de manera pública, y se ha utilizado en muchas otras aplicaciones y retos como base de extractor de características. Profunda y simple. Consiguió una tasa de error del 7.2%. En la Figura 8 se puede ver el dibujo de esta arquitectura.

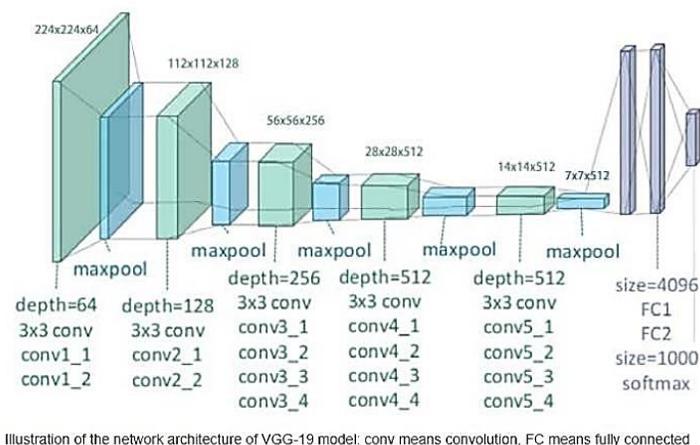


Fig. 9. Arquitectura VGGNet [19]

2.2.5 Arquitecturas para la clasificación de imágenes GoogLeNet.

Arquitectura de *GoogLeNet* (figura 9) fue la ganadora del reto en 2014. Por el nombre, puede saberse que es de Google, y sus autores hacen un pequeño tributo al autor de la red *LeNet*, ya que parte de su trabajo se basó en esa red. A pesar de ello, esta red tiene bastantes diferencias con sus predecesoras Dejó el error en 6.7%.

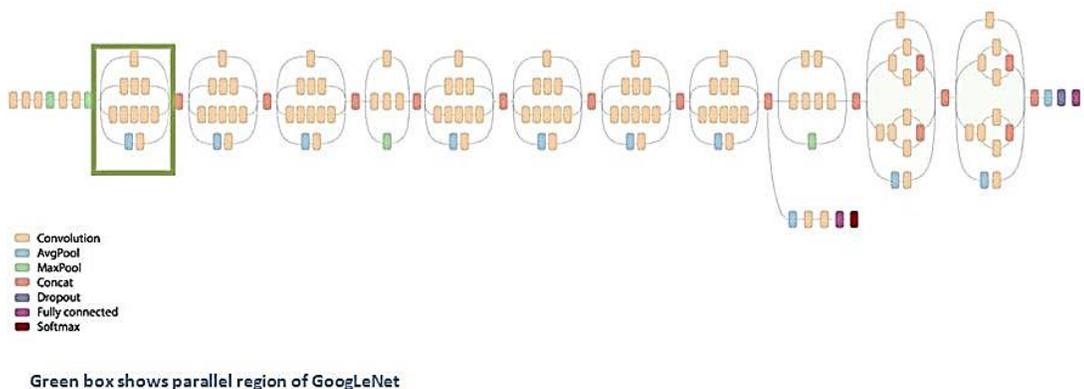


Fig. 10. Arquitectura de red neuronal GoogLeNet [19].

Si *VGG* destacó por su simpleza, *GoogLeNet* destaca, precisamente, por lo contrario. Esta red fue la primera red que se apartó del enfoque general que había hasta el momento, de simplemente apilar capas de redes convolucionales agrupadas una encima de la otra de manera secuencial.

2.2.6 Arquitecturas para la clasificación de imágenes ResNet

En 2015 Microsoft presentó una CNN basada en bloques residuales, *ResNet*. Fue ganadora del reto ya que bajó la tasa de error hasta el 3.57%. Consiguió con los bloques residuales y las 152 capas tener menos complejidad al entrenar que *VGGNet*. Con esta nueva arquitectura se batieron los récords en clasificación, localización y detección de objetos. Además, marcó un hito, ya que los errores de las personas en tareas de visión artificial oscilan entre un 5% y un 10%, dependiendo de las habilidades y experiencias de cada persona. Y esta red tiene una precisión mayor. En la siguiente Figura 10 se puede ver su arquitectura:

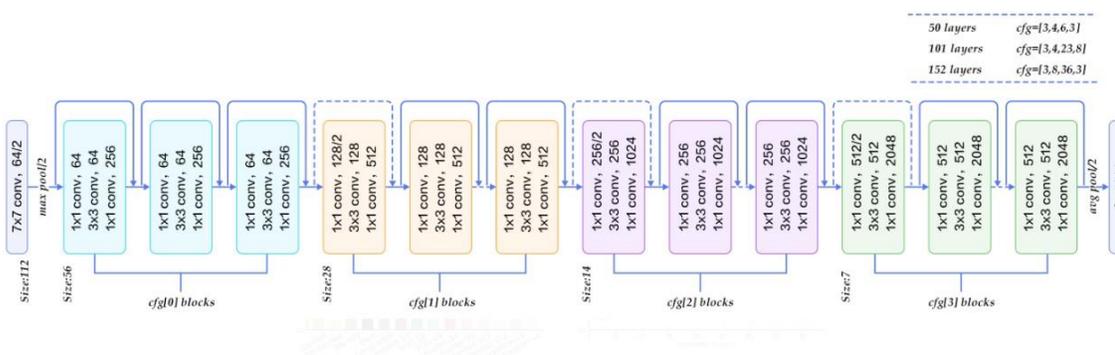


Fig. 11 ResNet bloques residuales [19].

2.2.7 Arquitecturas para la clasificación de imágenes ResNeXt

La arquitectura finalista del reto en 2016 es una ampliación del modelo *ResNet*, ya que añade una dimensión más, la cardinalidad a la arquitectura *ResNet*, quedando los bloques de la siguiente manera en la figura 11.

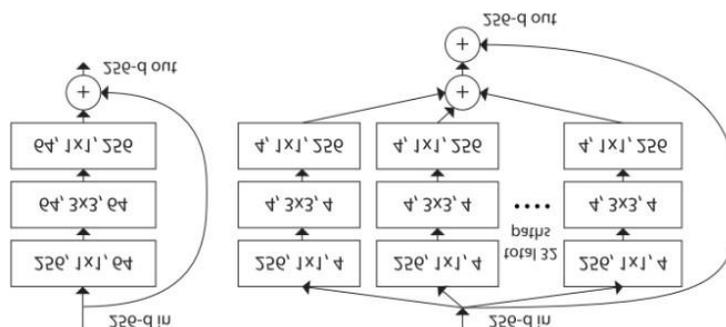


Fig. 12 Arquitecturas ResNeXt ampliación de Resnet [19].

2.2.8 Arquitecturas para la clasificación de imágenes Trimps-Soushen

La arquitectura ganadora la presentó *Trimps- Soushen*. *Trimps* es el instituto de investigación de tecnología para la seguridad china y *Soushen* traducido del chino significa “Dios de la búsqueda”. La arquitectura que presentó, y con la cual dejó el error por debajo del 3%, no presentaba ninguna tecnológica innovadora ni ninguna novedad. De hecho, no presentaron ningún artículo científico. Simplemente expusieron su trabajo en el *European Conference on Computer Vision*. Usaron modelos pre-entrenados con *Inception-v3* (40), *Inception- v4*, Pre-activation *ResNet-200* y *Wide ResNet* fusionando los resultados de los modelos. Observaron los resultados y se centraron en las diez categorías donde hubo más errores. Observando la Figura 11 se puede ver que no hay ningún modelo que supere a otro en ninguna categoría con respecto a los errores [19].

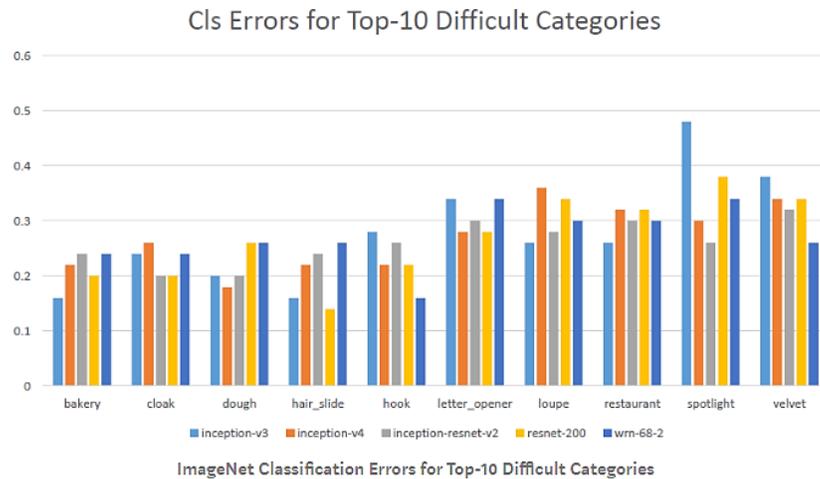


Fig. 13 Arquitectura categoría con respecto a los errores [19].

2.2.9 Arquitecturas para la clasificación de imágenes Inception V3

Inception busca disminuir el cuello de botella. En esta versión se enfoca que las redes neuronales trabajen óptimamente cuando las convoluciones no hacen redimensiones a las imágenes, ya que al eliminar fragmentos de estas, puede generar una pérdida de información, y a esta consecuencia se la conoce como cuello de botella. Esta versión número 3 de arquitecturas de redes neuronales profundas desarrolladas por Google llamadas Inception, Esta fue pre-entrenada usando el dataset compuesto por 1000 clases ImageNet y consta de las siguientes 13 capas [20]:

- a) Capa convolucional (299x299x3).
- b) Capa convolucional (149x149x32).
- c) Capa convolucional con *padding* (147x147x32).
- d) *Pool* (147x147x64).
- e) Capa convolucional (73x73x64).
- f) Capa convolucional (71x71x80).
- g) Capa convolucional (35x35x192).
- h) Capa 3xInception (35x35x288).

- i) Capa 5xInception (17x17x768).
- j) Capa 2xInception (8x8x1280).
- k) *Pool* (8x8x2048).
- l) Capa linear (1x1x2048).
- m) Capa SoftMax (1x1x1000) [20].

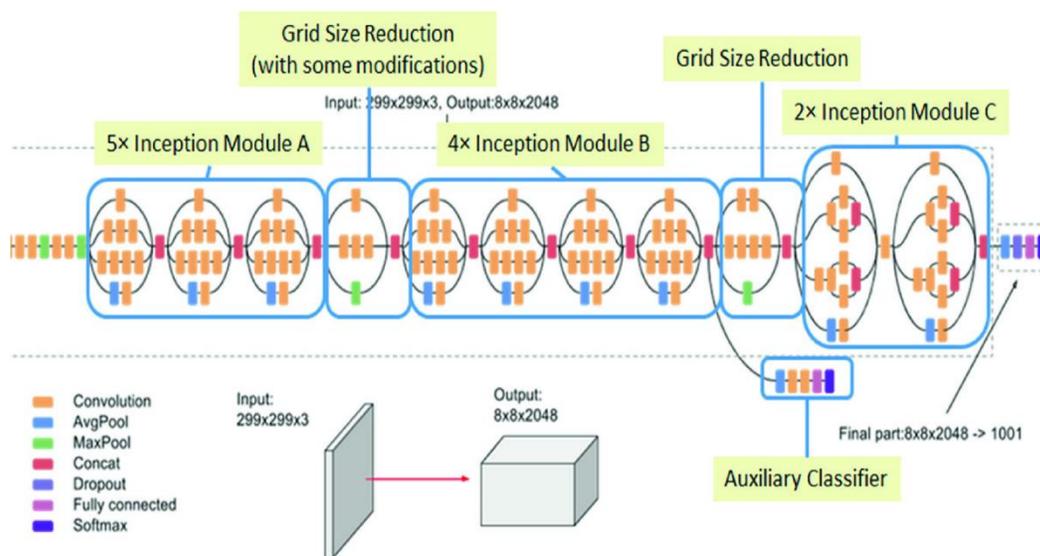


Fig. 14 Arquitectura Inception. [20]

2.3 Estado del arte

A continuación, se presentan los resultados donde se usó el repositorio scopus para obtener los datos e información de los artículos relacionados con palabras claves como: face shape, glasses and frames con el objetivo que se le da al desarrollo de este tipo de proyectos, a continuación, se presentan los artículos más relevantes que hasta la fecha han sido encontrados en las búsquedas realizadas. El primer trabajo que se encontró en la búsqueda corresponde al “An Interview Study

on Children's Spectacle Frame Fit" [25]. El cual hace referencia a estudios han examinado problemas con el ajuste de la montura de las gafas, y el diseño de marcos apropiado para niños. Donde el autor buscaba identificar prácticas y problemas en esta área, para lo cual se llevaron a cabo entrevistas semiestructuradas en Hong Kong con los ópticos dispensadores, los niños que usan gafas y los padres de los niños.

Este estudio proporcionó un análisis de datos que mostró el ancho de la montura, las almohadillas nasales entre otros aspectos. Pero el estudio concluyo que no existe un estándar de ajuste entre caras y marcos, y el ancho de la sien fue la referencia principal utilizada por los ópticos dispensadores para ayudar los niños eligen monturas de gafas, este estudio llego a la conclusión de que los ópticos dispensadores son importantes actores importantes en la selección de monturas de gafas más apropiadas para los niños, pero es posible que no puedan resolver los problemas de ajuste en el área de la nariz y las orejas debido a la deficiencia en el diseño de marcos. Entre los estudios analizados este en particular habla sobre los problemas en diseño y ajuste de gafas para niños, estudios adicionales sobre el diseño de monturas de gafas debe investigar los rasgos faciales y las necesidades especiales de los niños.

El estudio de Stephen Young, Friska Natalia, Sud Sudirman and Chang Seong Ko titulado "Eyeglasses frame selection based on oval face shape using convolutional neural network" [7]. Que en español es "selección de marco de anteojos según la forma de la cara ovalada uso de la red neural convolucional" este estudio permitió *analizar el método que se puede utilizar es la red neuronal convolucional*. En este caso, por utilizando la red neuronal convolucional, identificará que la forma de la cara de una persona es ovalada o no. Este método se implementó en el sistema operativo Android y se utilizará el lenguaje Python para ayudar en el tratamiento de los datos que recuperamos y observamos del motor de búsqueda de Google. Viola-Jones El algoritmo es útil para analizar la existencia de caras en el procesamiento de imágenes. La salida que se logrará es una aplicación que puede dar al usuario una opción de marcos de gafas basadas en formas faciales.

A través de esta aplicación, el usuario puede probar una variedad de marcos. Con práctica en lugar de tener que probarlo en forma física uno por uno.

Ahora bien, el estudio titulado, “Probabilistic neural network and invariant moments for men face shape Classification” del autor Romi Fadillah Rahmat, Muhammad Dian Syahputra, Ulfi La clasificación de la forma de la cara es útil para arreglar los marcos de las gafas según las personalidades, como la selección del corte de pelo, la selección del maquillaje facial, la selección de gafas marcos o incluso la selección de camisetas adecuadas. Es por ello por lo que los autores analizaron la forma de la cara en los hombres es dividido en seis formas, a saber: ovalada, redonda, diamante, rectángulo, triángulo y cuadrado [26]. Donde se puede evidenciar según el estudio que la determinación de la forma facial ha sido introducida por muchos expertos en belleza, pero para la sociedad, en general, todavía es un poco difícil clasificarlo porque la forma de cada cara es casi igual y la medición manual requiere un largo proceso. Por eso necesita un método para clasificar la forma del rostro de forma rápida y precisa. Al método propuesto en esta investigación es Red neuronal de probabilidad e Invariante, Donde las imágenes de rostros de hombres se utilizan como entrada para el procesamiento de imágenes. Donde los investigadores categorizaron las etapas antes de la clasificación son el preprocesamiento de imágenes (escala de grises, escala y Gabor Filtrar). Luego, caracterice la extracción usando Invariant Moments. El paso final es clasificación utilizando Probability Neural Network. Después de realizar la prueba a 90 datos entrenamiento y 30 pruebas de datos, se concluyó que el método propuesto tiene la capacidad para clasificar la forma de la cara de los hombres con una precisión del 80%. Lo que permite tomar como referencia este antecedente como fundamento para la presente investigación.

La clasificación de la forma de la cara es un proceso vital para elegir pestañas, peinado y maquillaje facial adecuados, y selección de monturas de gafas adecuadas de acuerdo con las pautas de expertos. Esta investigación permitió analizar la medición de las características del rostro por expertos en belleza manualmente cuesta tiempo y esfuerzos. Sin embargo, los métodos existentes

abordan muchos desafíos debido a la complejidad de la geometría de la cara y variación de sus características. Donde los autores proponen el método que se basa en la fusión de las características aprendidas por CNN con características hechas a mano representadas por histograma de gradientes orientados (HOG) y puntos de referencia faciales ha demostrado ser eficiente en la identificación de la forma facial. El resultado obtenido del estudio demostró que el método propuesto es prometedor para identificar la forma de la cara logra una precisión del 81,1% [27]. En este contexto un cabello es la característica más singular que puede mejorar los rasgos faciales de una persona en especial las mujeres por ello el estudio se basa en cuál es el corte de cabello adecuado según la forma de la cara de la mujer, lo cual da algunas pautas de cómo podrían llegar a ser los marcos de gafas según el corte de cabello. Donde los investigadores crearon un modelo que tiene la capacidad de clasificar la forma de la cara, cuando un usuario subió un retrato de sí misma, creando a su vez un aprendizaje automático, Se utilizaron bibliotecas para identificar los puntos de referencia de la imagen de la cara, con un algoritmo de clasificación ha utilizado para recomendar la mayoría peinado o corte de pelo adecuado según el rostro detectado forma, longitud del cabello e información recopilada del cabello expertos [28].

Por otro lado, el estudio de Adonis Emmanuel de la universidad de Philippines Diliman, en su investigación titulada “Face shape classification using Inception v3” [28]. Presentan los resultados experimentales obtenidos al reentrenar la última capa del modelo Inception v3 en la clasificación de imágenes de rostros humanos en una de las cinco formas básicas de rostros. La precisión del modelo Inception v3 reentrenado se comparó con la de los siguientes métodos de clasificación que utilizan las relaciones y ángulos de distancia de los puntos de referencia faciales como características: análisis discriminante lineal (LDA), máquinas de vectores de soporte con kernel lineal (SVM-LIN), máquinas de vectores de soporte con núcleo de función de base radial (SVM-RBF), redes neuronales artificiales o perceptrón multicapa (MLP) y k vecinos más cercanos (KNN). Todos los clasificadores fueron entrenados y probados utilizando un total de 500 imágenes de celebridades femeninas con formas de rostro conocidas recopiladas de Internet.

Los resultados muestran que la precisión del entrenamiento y la precisión general varían del 98.0% al 100% y del 84.4% al 84.8% para Inception v3 y del 50.6% al 73.0% y del 36.4% al 64.6% para los otros clasificadores, dependiendo del tamaño del conjunto de entrenamiento utilizado. Este resultado muestra que el modelo Inception v3 reentrenado pudo ajustar bien los datos de entrenamiento y superar a los otros clasificadores sin la necesidad de seleccionar características específicas para incluirlas en el entrenamiento del modelo.

Esta investigación se diferencia de las aquí citadas dado a la funcionalidad que se pretende con la creación de un sistema recomendador de marcos para lentes que pueda ser utilizado como una estrategia para apoyar a las empresas e instituciones encargadas de proporcionar los lentes oftálmicos, y así construir una herramienta que genere un desarrollo para aportar a mejorar la calidad de vida de las personas incentivando la adopción de marcos para gafas por cada paciente, el EYEWEAR CONSULTANT se diferencia del proyecto “Eyeglasses frame selection based on oval face shape using convolutional neural network” ya que se entrena la red neuronal con 5 tipos de rostros diferentes: ovalado, redondo, cuadrado, rectangular y tipo corazón permitiendo al sistema realizar una recomendación diferente para cada tipo de característica facial, ofreciendo los marcos que se encuentren disponibles en el inventario ofertados por cada tienda óptica.

Como antecedentes se mencionan: Yenebez Moreno en el 2011 presentó un estudio en Madrid España, el cual mostró como objetivo poner en marcha la plataforma web Gafasnet.com para la comercialización de productos ópticos tales como gafas de sol, gafas graduadas, gafas especiales para una actividad determinada como deporte, además de lentes de contacto, cosmética ocular, parafarmacia ocular, en la empresa Investigación y Distribuidora de Productos Ópticos Internet S.L., para ello como herramientas hizo uso de un software libre haciendo especial énfasis en las plataformas Prestashop y Wordpress, concluyendo que en este proyecto se ha tenido la oportunidad de constatar que las principales plataformas de software libre existentes no están preparadas para comercializar de forma cómoda un producto óptico graduado. La puesta en marcha

de este proyecto ha significado hacer enormes modificaciones a la solución software elegida, Prestashop, hasta el punto de resultar dudosa la viabilidad de actualizar la versión de esa parte del software, por ello se recomendó hacer mejoras en futuras investigaciones que tomen como modelo este antecedente.

Asimismo, Espinoza Moncayo & Peña Mendoza en el 2015 realizaron una investigación en Guayaquil la cual presentó como objetivo elaborar un diseño e implementar un prototipo de gafas electrónicas que ayude dando una alerta a las personas no videntes, indicando la distancia a la que se encuentran un obstáculo mediante un celular, de tal manera que este pueda evadirlo al momento de movilizarse, utilizó como herramientas unas tarjetas de control, antenas, módulos de comunicación inalámbrica, sensores y software como son: PIC KIT 2, MICROBASIC PRO, APP INVENTOR 2, seguidamente se procedió al montaje de las tarjetas de control y los sensores en gafas para finalmente realizar las pruebas necesarias y verificar el correcto funcionamiento del proyecto, con los resultados obtenidos se presenta un prototipo de gafas electrónicas que se comunican vía bluetooth a un celular capaz de advertir tanto en frente como a los lados la presencia de un obstáculo.

Igualmente, Palacio Méndez en el 2015, mostró en Nicaragua una investigación que tuvo como objetivo analizar cómo son adaptadas las lentes oftálmicas en las ópticas de Nicaragua, septiembre - noviembre 2015. Para conocer esta realidad se ha realizado un estudio descriptivo, transversal, cuyo universo son todos los pacientes portadores de lentes, del cual se tomará una muestra de 300 personas, siendo los criterios de inclusión todos los pacientes portadores de lentes y los criterios de exclusión no portadores. La recolección de datos ha sido la mediante aplicación de encuestas. El procesamiento de la Información se realizó a través de Microsoft Excel, mediante la elaboración de tablas y gráficos que demuestran la estimación de las variables. Como conclusiones logró alcanzar los objetivos planteados para la investigación y a pesar de que en las ópticas de Nicaragua no se realizan las adaptaciones de acuerdo con el protocolo una buena parte

de la población usuaria de lentes oftálmicos logra adaptarse y refieren buena calidad visual y satisfacción.

CAPITULO III METODOLOGÍA

Esta investigación fue desarrollada bajo la metodología del Patrón de Investigación Iterativo, se definieron los 3 ciclos relacionados con los objetivos propuestos cada uno conformado de 4 fases iterativas relacionadas con: observación, identificación, desarrollo y pruebas, donde se generarán los diferentes entregables y productos del proyecto. El proyecto tuvo una etapa inicial que se realizó de forma constante en todo el tiempo de duración de la investigación relacionada con la actualización del estado del arte y la redacción de documentos e informes técnicos [31].

Ciclo 1: Se definió un conjunto y tipo de datos establecido en el estado del arte del proyecto relacionado con la clasificación de formas de rostros de personas según su morfología y técnicas de visagismo para ser utilizado en algoritmos de clasificación.

- **Observación:** Se recolectó la información y conjuntos de datos donde se encontraron variedad de rostros humanos en internet y en Dataset validados.
- **Identificación:** Se seleccionó la información que permitió la clasificación de rostros basados en la morfología y técnicas del visagismo.
- **Desarrollo:** Se descargaron imágenes de rostros de personas donde se pueden observar y demarcar las regiones de interés.

- **Pruebas:** Se generó varios archivos donde se mostró el rostro humano, donde permitió su identificación y clasificación de formas de rostros para entrenar el algoritmo planteado.

Ciclo 2: Se construyó un modelo de aprendizaje profundo basado en una red neuronal convolucional para la clasificación de tipos de rostros donde permitió generar recomendaciones hacia los usuarios relacionado con marcos para lentes.

- **Observación:** Se determinó las arquitecturas de redes neuronales convolucionales, que permitieron un aprendizaje profundo para clasificación correcta de la forma de los rostros.
- **Identificación:** Se identificó la arquitectura y el algoritmo que permitió obtener buenos resultados para el entrenamiento y la interpretabilidad obteniendo como resultado la detección y clasificación de rostros utilizando imágenes de dataset validados.
- **Desarrollo:** Se construyeron herramientas que permiten un buen entrenamiento del sistema para la validación y prueba de la detección y clasificación de rostros.
- **Pruebas:** Se probó el algoritmo entrenado de redes neuronales convolucionales para toma de decisiones de recomendaciones para marcos de lentes.

Ciclo 3: Se evaluó el modelo de clasificación de tipos de rostros utilizando métricas de validación como la matriz de confusión, coeficiente Kappa y porcentajes de error.

- **Observación:** Se organizó las imágenes que contienen tipos de rostros cargados desde el dataset validado en el estado del arte.
- **Identificación:** Se identificaron las métricas de validación como la matriz de confusión, coeficiente kappa y porcentajes de error para una correcta clasificación de rostros.

- **Desarrollo:** Se ejecutó el algoritmo basado en el aprendizaje profundo para la clasificación de los rostros y así evaluar la precisión del sistema.
- **Pruebas:** Se hizo ejecución del sistema de clasificación de rostros con imágenes del dataset validado y calificar el tipo de recomendación para los marcos de gafas.

TABLA. I Actividades y resultados relacionados con el desarrollo del proyecto.

Objetivo	Actividad	Instrumento	Producto esperado
Definir un conjunto de datos establecido en el estado del arte del proyecto relacionado con la clasificación de formas de rostros de personas según su morfología y técnicas de visagismo para ser utilizado en algoritmos de clasificación.	Seleccionar un conjunto de datos relacionado con tipos de rostro de personas, el dataset debe estar validado por la comunidad científica.	Computador, fuentes bibliográficas.	Dataset seleccionado y preprocesado para ser resuelto por una arquitectura de red neuronal convolucional.
	Identificación de los diferentes tipos de rostros según conceptos de visagismo.	Computador, fuentes bibliográficas.	Informe técnico con las principales características de los distintos tipos de rostros.
Construir un modelo de aprendizaje profundo basado en una red neuronal convolucional para la clasificación de tipos de rostros que permita generar recomendaciones hacia los usuarios relacionado con marcos para lentes.	Implementar una arquitectura de red neuronal convolucional para clasificación de imágenes.	Computadores, Librerías: OpenCV, TensorFlow y lenguaje de programación Python	Informe técnico con la arquitectura implementada y los hiperparámetros ajustados.
	Construir un algoritmo para el entrenamiento de una red neuronal convolucional que facilite la clasificación de tipos de rostros	Computadores, Librerías: OpenCV, TensorFlow y lenguaje de programación Python	Código fuente del entrenamiento de la red neuronal.

Evaluar el modelo de clasificación de tipos de rostros utilizando métricas de validación como la matriz de confusión, coeficiente Kappa y porcentajes de error	Implementar el sistema recomendador utilizando librerías para la creación de interfaces gráficas: Tkinter.	Computadores, Sistema desarrollado.	Aplicación de escritorio con el sistema de recomendación de marcos de lentes.
	Aplicación de métricas relacionadas con la clasificación de imágenes.	Computadores, Sistema desarrollado.	Reporte con las métricas de evaluación del sistema.

3.1 Entrenamiento de la red neuronal

El entrenamiento de la red neuronal se trabajó con un conjunto de datos de 4000 imágenes divididas en 5 carpetas cada una con 800 tipos de rostros, llamadas de la siguiente manera: ovalado, rectangular, cuadrado, redondo y corazón. Con el conjunto de datos descargado se comenzó a entrenar la red neuronal utilizando la biblioteca KERAS que es una de las principales librerías de redes neuronales que permite crear de forma fácil modelos de deep learning. Admite múltiples motores de cálculos de redes neuronales, por tanto, se puede fusionar y ejecutar sobre tensorflow, Tensorflow, es una biblioteca de código abierto que facilita el proceso de entrenamiento de modelos de las redes neuronales convolucionales y aprendizaje automático

. En las primeras ejecuciones en el entrenamiento de la red neuronal con nuestro conjunto de datos tal cual se descargó no se tuvo buenos resultados ya que se estaba tomando como región de interés partes de la imagen que no se requerían donde los porcentajes de error y precisión no cambiaban un ejemplo es la imagen que se muestra en la figura 14 por lo tanto, se realizó la modificación del conjunto de datos, se evaluó y analizó cada imagen, al identificar que el espacio entre los bordes y la cabeza era una región irrelevante se decidió recortar de forma manual la imagen como se muestra en la figura 15, obteniendo como resultado que solo quedara el rostro mejorando los resultados de precisión.



Fig. 15. Ejemplo de imágenes descargadas del dataset de Kaggle [33]



Fig. 16. Ejemplo de imagen recortada manualmente [33]

Otras librerías que proporcionan funcionalidades y operaciones el cual nos permite implementar en nuestro modelo son

- NumPy es una biblioteca para el lenguaje de programación Python que da soporte para crear vectores y matrices grandes multidimensionales, junto con una gran colección de funciones matemáticas de alto nivel para operar con ellas.

Matplotlib es una biblioteca para la generación de gráficos a partir de datos contenidos en listas o arrays en el lenguaje de programación Python y su extensión matemática NumPy.

3.2 Código entrenamiento de RedNeuronal.

En las primeras líneas de código se importan las librerías que nos brindan ciertas operaciones y funcionalidades ya resultas por una compañía o desarrollador en particular

```
1 import matplotlib.pyplot as plt
2 import numpy as np
3 import os
4 import PIL
5 import tensorflow as tf
6 from tensorflow import keras
7 from tensorflow.keras import layers
8 from tensorflow.keras.models import Sequential
9 import pathlib
10 import cv2 as cv
```

Fig. 17 Importan las librerías

En las siguientes líneas de código se especifica la ruta donde se encuentran el dataset, se define la ruta de directorio y se encuentran las imágenes por su formato

```
12 path_dir='/content/drive/MyDrive/A-Grado/Dataset_recortado'  
13 data_dir = pathlib.Path(path_dir)  
14 image_count = len(list(data_dir.glob('*/*.jpg')))  
15 print(image_count)
```

Fig. 18 Se carga el data set.

Se definen los canales, el tamaño de las muestras, ancho, alto de las imágenes y el número de veces que se ejecuta el algoritmo

```
17 channels=3  
18 batch_size = 64  
19 img_height = 180  
20 img_width = 180  
21 epochs=80
```

Fig. 19 Parámetros iniciales del entrenamiento.

Se ejecuta la reprocesamiento de conjunto de datos el cual permite categorizar o clasificar las imágenes contenidas en el dataset

```

23 train_ds = tf.keras.preprocessing.image_dataset_from_directory(
24     data_dir,
25     validation_split=0.2,
26     subset="training",
27     seed=123,
28     image_size=(img_height, img_width),
29     batch_size=batch_size)
30
31 val_ds = tf.keras.preprocessing.image_dataset_from_directory(
32     data_dir,
33     validation_split=0.2,
34     subset="validation",
35     seed=123,
36     image_size=(img_height, img_width),
37     batch_size=batch_size)
38
39 class_names = train_ds.class_names
40 print(class_names)

```

Fig. 20 Convertir en formato dataset.

Permite normalizan los datos, permitiendo reemplazar 1-255 por 0-1

```

42 AUTOTUNE = tf.data.experimental.AUTOTUNE
43 train_ds = train_ds.cache().shuffle(1000).prefetch(buffer_size=AUTOTUNE)
44 val_ds = val_ds.cache().prefetch(buffer_size=AUTOTUNE)
45 normalization_layer = layers.experimental.preprocessing.Rescaling(1./255)
46 num_classes = len(class_names)

```

Fig. 21 Se modifica el dataset.

Se define una capa secuencial que permite que las capas se agreguen en secuencia

```

48 data_augmentation = keras.Sequential(
49     [
50         layers.experimental.preprocessing.RandomFlip("horizontal",
51                                                     input_shape=(img_height,
52                                                         img_width,
53                                                         3)),
54         layers.experimental.preprocessing.RandomRotation(0.1),
55         layers.experimental.preprocessing.RandomZoom(0.1)
56     ]
57 )

```

Fig. 22 Aumento de datos del dataset.

Se crean los modelos iniciales, utilizamos la función de activar ReLU para que active la neurona cuando es mayor a 0

```

59 model = Sequential([
60     #1st CNN layer
61     layers.Conv2D(64, 3, padding='same', activation='relu', input_shape=(img_height, img_width, channels)),
62     layers.BatchNormalization(),
63     layers.MaxPooling2D(pool_size = (2,2)),
64     layers.Dropout(0.30),
65     #2nd CNN layer
66     layers.Conv2D(256, 5, padding='same', activation='relu'),
67     layers.BatchNormalization(),
68     layers.MaxPooling2D(pool_size = (2,2)),
69     layers.Dropout(0.30),
70     #3rd CNN layer
71     layers.Conv2D(512, 3, padding='same', activation='relu'),
72     layers.BatchNormalization(),
73     layers.MaxPooling2D(pool_size = (2,2)),
74     layers.Dropout(0.30),
75     #4th CNN layer
76     layers.Conv2D(512, 3, padding='same', activation='relu'),
77     layers.BatchNormalization(),
78     layers.MaxPooling2D(pool_size = (2,2)),
79     layers.Dropout(0.30),
80     layers.Flatten(),
81     #Fully connected 1st layer
82     layers.Dense(512, activation='relu'),
83     layers.BatchNormalization(),
84     layers.Dropout(0.30),
85     #Fully connected 2nd layer
86     layers.Dense(512, activation='relu'),
87     layers.BatchNormalization(),
88     layers.Dropout(0.30),
89     layers.Dense(num_classes, activation='softmax')
90 ])

```

Fig. 23 Se crea la arquitectura de la red neuronal convolucional.

Se configura y se define donde guardar el modelo entrenado


```
120 epochs_range = range(epochs)
121
122 plt.figure(figsize=(8, 8))
123 plt.subplot(1, 2, 1)
124 plt.plot(epochs_range, acc, label='Training Accuracy')
125 plt.plot(epochs_range, val_acc, label='Validation Accuracy')
126 plt.legend(loc='lower right')
127 plt.title('Training and Validation Accuracy')
128
129 plt.subplot(1, 2, 2)
130 plt.plot(epochs_range, loss, label='Training Loss')
131 plt.plot(epochs_range, val_loss, label='Validation Loss')
132 plt.legend(loc='upper right')
133 plt.title('Training and Validation Loss')
134 plt.show()
135
136 test_loss, test_acc = model.evaluate(val_ds, verbose=2)
137 print(test_loss, test_acc)
```

Fig. 28 Generación de métricas visuales

CAPÍTULO IV RESULTADOS Y ANÁLISIS

4.1 Interfaz gráfica

Para esta parte del código se trabajó con tkinter que es una librería que permite implementar la interfaz de nuestro proyecto en el lenguaje de programación Python. A continuación, se explicará paso a paso el código.

- a) Cuando se abre el sistema recomendador de gafas inicia en la captura de la foto al usuario como se ve en la figura 29.

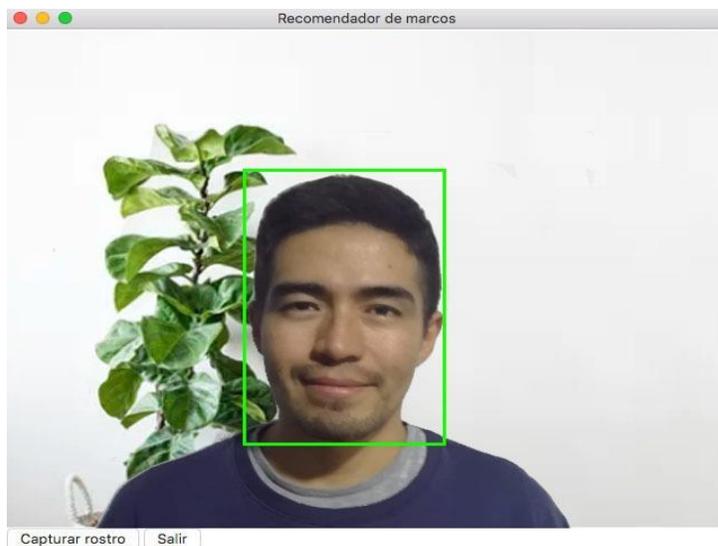


Fig. 29 Inicia el sistema en la captura de la imagen: Fuente propia.

- b) Figura 30 la cual muestra cómo se procesa la imagen sacando la región de interés y se da la opción de ver la recomendación de marcos o de tomar un nuevo rostro.



Fig. 30 Rostro recortado: Fuente propia.

- c) Da la recomendación de los marcos según la forma del rostro, como se muestra en la figura 31.



Fig. 31 Recomendación de marcos: Fuente propia.

4.1.1 Evidencia del código y explicación.

```

45     def __del__(self):
46         print("Clase Videocapture-----del")
47         if self.vid.isOpened():
48             self.vid.release()
49             self.out.release()
50             cv2.destroyAllWindows()

```

Fig. 32 Se crea el canvas y los botones de tomar fotografía y salir: Fuente propia

De la línea 32 a la línea 51 se crea el canvas o el lienzo donde se muestra lo que se enfoca con la cámara, así como también el menú con los botones de tomar fotografía, botón de salir, tal como se muestra en la Figura 29.

```

52     def tick(self):
53         print("Clase Videocapture-----tick")
54         self.now = dt.datetime(1, 1, 1).now()
55         self.elapsedTime = self.now - self.zeroTime
56         self.time2 = self.elapsedTime.strftime('%H:%M:%S')
57         if self.time2 != self.lastTime:
58             self.lastTime = self.time2
59             self.T.config(text=self.time2)
60         self.updwin=self.T.after(100, self.tick)
61

```

Fig.33. Función capturar fotografía: Fuente propia

De la línea 54 a la línea 69 se implementa la función de capturar fotografía, permite guardar una fotografía a partir de la cámara permite cargar o hacer lectura de la imagen capturada, como también recortar la imagen de acuerdo con el marco guía.

Dentro de la misma función desde la línea 70 a la línea 77 se genera un popup o ventana emergente donde se muestra la imagen recortada con un menú con botones de ver recomendación y nuevo rostro, tal como se muestra en la Figura 30.

```

81     def VerRecomendación(self, tipoRostro):
82         rostro=tipoRostro
83
84         if rostro=="Redondo":
85             imgText = cv2.imread('RostroCuadrado/Redondo.png')
86             img1 = cv2.imread('RostroCuadrado/Redondo0.png')
87             img2 = cv2.imread('RostroCuadrado/Redondo1.png')
88             img3 = cv2.imread('RostroCuadrado/Redondo2.png')
89             img4 = cv2.imread('RostroCuadrado/Redondo3.png')
90             img5 = cv2.imread('RostroCuadrado/Redondo4.png')
91
92         elif rostro=="Agatado":
93             imgText = cv2.imread('RostroOvalado/Agatado.png')
94             img1 = cv2.imread('RostroOvalado/Agatado0.png')
95             img2 = cv2.imread('RostroOvalado/Agatado1.png')
96             img3 = cv2.imread('RostroOvalado/Agatado2.png')
97             img4 = cv2.imread('RostroOvalado/Agatado3.png')
98             img5 = cv2.imread('RostroOvalado/Agatado4.png')
99
100        elif rostro=="Ovalado":
101            imgText = cv2.imread('RostroRectangular/Ovalado.png')
102            img1 = cv2.imread('RostroRectangular/Ovalado0.png')
103            img2 = cv2.imread('RostroRectangular/Ovalado1.png')
104            img3 = cv2.imread('RostroRectangular/Ovalado2.png')
105            img4 = cv2.imread('RostroRectangular/Ovalado3.png')
106            img5 = cv2.imread('RostroRectangular/Ovalado4.png')
107

```

Fig. 34 Define marcos a mostrar: Fuente propia.

De la línea 85 a la línea 125 se implementa la función de Ver recomendación, dependiendo del String que se recibe desde nuestra red neuronal se define qué tipos de marcos mostrar.

```

129 def vconcat_resize(img_list, interpolation
130                      = cv2.INTER_CUBIC):
131     w_min = min(img.shape[1]
132                for img in img_list)
133
134     im_list_resize = [cv2.resize(img,
135                                (w_min, int(img.shape[0] * w_min / img.shape[1])),
136                                interpolation = interpolation)
137                      for img in img_list]
138     return cv2.vconcat(im_list_resize)
139 def hconcat_resize(img_list,
140                   interpolation
141                   = cv2.INTER_CUBIC):
142     h_min = min(img.shape[0]
143                for img in img_list)
144     im_list_resize = [cv2.resize(img,
145                                (int(img.shape[1] * h_min / img.shape[0]),
146                                h_min), interpolation
147                                = interpolation)
148                      for img in img_list]
149     return cv2.hconcat(im_list_resize)
150 def concat_tile_resize(list_2d,
151                       interpolation = cv2.INTER_CUBIC):
152     img_list_v = [hconcat_resize(list_h,
153                                interpolation = cv2.INTER_CUBIC)
154                  for list_h in list_2d]
155     return vconcat_resize(img_list_v, interpolation=cv2.INTER_CUBIC)
156 im_tile_resize = concat_tile_resize([[imgText],
157                                     [img1, img2,
158                                     img3],
159                                     [img4, img5]])
160
161 cv2.imshow('Recomendador', im_tile_resize)
162 cv2.waitKey(0)

```

Fig. 35 Muestra los marcos recomendados: Fuente propia

De la línea 129 a 162 se crea una matriz en interfaz para mostrar los marcos recomendados como se muestra en la figura 31.

```

7 class VideoCapture:
8     def __init__(self, video_source=0):
9
10        self.vid = cv2.VideoCapture(video_source)
11        if not self.vid.isOpened():
12            raise ValueError("Unable to open video source", video_source)
13
14        args=CommandLineParser().args
15
16        STD_DIMENSIONS = {
17            '480p': (640, 480),
18            '720p': (1280, 720),
19            '1080p': (1920, 1080),
20            '4k': (3840, 2160),
21        }
22        res=STD_DIMENSIONS[args.res[0]]
23        print(args.name,self.fourcc,res)
24        self.out = cv2.VideoWriter(args.name[0]+'.'+args.type[0],self.fourcc,10,res)
25        self.vid.set(3,res[0])
26        self.vid.set(4,res[1])
27
28        self.width,self.height=res
29

```

Fig. 36 Muestra los frames: Fuente propia

De la línea 8 a 28 de la clase VideoCapture se define las dimensiones en las que se muestra los frames o lo que se captura con la cámara.

```

30 def get_frame(self):
31     if self.vid.isOpened():
32         ret, frame = self.vid.read()
33         if ret:
34             starting_point = (200, 100)
35             ending_point = (400, 400)
36             color= (0, 255, 0)
37             thickness = 2
38             cv2.rectangle(frame, starting_point, ending_point, color, thickness)
39             return (ret, cv2.cvtColor(frame, cv2.COLOR_RGB2BGR))
40         else:
41             return (ret, None)
42     else:
43         return (ret, None)
44

```

Fig. 37 Validar cámara encendida: Fuente propia

De la línea 30 a 43 de la clase (VideoCapture) se implementa la función que permite validar si efectivamente la cámara función y/o está encendida.

```

45     def __del__(self):
46         print("Clase Videocapture-----del")
47         if self.vid.isOpened():
48             self.vid.release()
49             self.out.release()
50             cv2.destroyAllWindows()

```

Fig. 38 Destruir ventanal salir: Fuente propia

De la línea 45 a 50 de la clase (videoCapture) se implementa la función para destruir la ventana al salir.

```

52     def tick(self):
53         print("Clase Videocapture-----tick")
54         self.now = dt.datetime(1, 1, 1).now()
55         self.elapsedTime = self.now - self.zeroTime
56         self.time2 = self.elapsedTime.strftime('%H:%M:%S')
57         if self.time2 != self.lastTime:
58             self.lastTime = self.time2
59             self.T.config(text=self.time2)
60             self.updwin=self.T.after(100, self.tick)
61

```

Fig. 39 Capturar hora y fecha: Fuente propia

De la línea 52 a 60 de la clase (videoCapture) se implementa la función que permite capturar la hora y fecha del sistema (Estos datos se usan como parámetro en el nombre de la fotografía del cliente).

El mejor modelo de clasificación de red neuronal artificial convolucional obtuvo precisiones generales en el entrenamiento y validación de 73.8% y 53.3% respectivamente. La precisión más baja de una clase corresponde a los rostros ovalados con un 69.3% y 45.5% en las etapas de entrenamiento y validación. La clase con mayor precisión que se obtuvo después del entrenamiento del modelo corresponde a los tipos de rostros oblongos con valores de 80.9%, sin embargo, en la etapa de validación la clase con mayor precisión pertenece a los tipos de rostro redondos con un 60.1%. En la tabla 1 y 2, se presentan las demás métricas de evaluación que se tuvieron en cuenta durante el proceso de entrenamiento y validación.

Tabla 2. Métricas de evaluación obtenidas durante el proceso de entrenamiento del modelo de red neuronal convolucional (2801 imágenes de entrenamiento).

Class	Precision	Recall	F1	AP	ROC AUC
Square	73.3%	73.4%	73.4%	80.2%	92.8%
Round	74.5%	77.5%	76.0%	82.1%	93.8%
Oval	69.3%	67.4%	68.3%	75.1%	90.0%
Oblong	81.7%	74.4%	77.9%	85.6%	94.8%
Heart	70.2%	78.0%	73.9%	83.1%	93.5%

Tabla 3. Métricas de evaluación obtenidas durante el proceso de validación del modelo de red neuronal convolucional (1199 imágenes de prueba).

Class	Precision	Recall	F1	AP	ROC AUC
Square	51.3%	47.9%	49.6%	50.7%	79.1%
Round	60.1%	55.8%	57.9%	58.5%	82.8%
Oval	45.5%	35.8%	40.1%	41.7%	72.8%
Oblong	57.7%	56.5%	57.1%	56.8%	83.6%
Heart	51.8%	53.3%	52.6%	58.1%	81.1%

4.2. Conclusiones.

Teniendo en cuenta el bajo porcentaje de precisión de entrenamiento con las imágenes que conforman el conjunto de datos se determina que el modelo está tomando toda la información de la imagen, es decir que además del rostro de la persona también guarda información como el fondo de la imagen y otras partes del cuerpo. Es por esta razón que las imágenes fueron modificadas a través del ajuste de una región de interés que abarca el alto y el ancho del rostro. El objetivo de este ajuste está relacionado con la disminución del ruido generado debido al ambiente natural que presentan las imágenes, esta limitación de espacio permite al modelo de red neuronal enfocarse en el área de interés para el proyecto.

Se evidencia que a mayor número de datos para entrenar el modelo, el porcentaje de precisión aumenta, teniendo en cuenta esta información el sistema de recomendación de marcos para gafas fue construido con un conjunto de datos validado que consta de 4.000 imágenes de los cuales se incluyen los 5 tipos de rostro trabajados.

Python cuenta varias librerías que permiten la optimización de trabajo y rendimiento, es por ello que una de las librerías que nos brinda la facilidad de implementar una interfaz gráfica y crear aplicaciones de escritorio es Tkinter. De esta manera que el sistema eyewear consultant está basado en una aplicación de escritorio utilizando la librería en mención, el cual puede ser instalada en equipos que tengan buena capacidad de memoria para tener un funcionamiento óptimo.

La precisión del modelo entrenado es importante para identificar la confianza de la clasificación. Una de las mejores formas de presentar los resultados de estas predicciones es la matriz de confusión, lo cual es una de las métricas más intuitivas y simples para encontrar la exactitud del modelo, Se identifica cada tipo de rostro, sus predicciones correctas y sus

predicciones incorrectas que hizo el modelo, es decir que tan efectivo fue al identificar un tipo de rostro. Como resultado general se obtuvo una precisión del modelo de 73,4% el cual resulta de la operación 2955 , que es número de predicciones correctas dividido 4005 que es el total de imágenes clasificadas en el conjunto de datos.

4.3 Trabajos futuros.

Como trabajo a futuro se plantea enriquecer el conjunto de datos trabajado con rostros masculinos disminuyendo el ruido de las imágenes buscando un fondo más limpio para estas y así obtener mejores resultados en el entrenamiento a futuro.

Implementar en el sistema eyewear consultant la opción de medirse los marcos existentes en las ópticas según la recomendación dada trabajando con realidad aumentada.

Trabajar en una aplicación móvil del sistema recomendador EYEWEAR CONSULTANT donde se haga un cobro único con fines comerciales para que las instituciones o tiendas ópticas tengan fácil acceso a él para que lo puedan implementar en sus negocios con disponibilidad en tiendas de sistema Android y IOS.

Combinar las variables de las técnicas del visagismo con el modelo entrenado para mejorar las precisiones en las métricas que validen el funcionamiento óptimo del EYEWEAR CONSULTANT

Referencias

- 1] Organizacion Mundial de la Salud, «La OMS presenta el primer informe mundial sobre la vision,» 2019. [En línea]. Available: <https://www.who.int/es/news/item/08-10-2019-who-launches-first-world-report-on-vision> .

- 2] Organizacion Mundial de la Salud, «Una mejor atencion oftalmica para todos que puede reportar beneficios economicos,» 2012. [En línea]. Available: https://www.who.int/mediacentre/news/notes/2012/eye_care_20121004/es/.

- 3] Ministerio de salud, «Veo bien aprendi bien,» 2021. [En línea]. Available: <https://www.minsalud.gov.co/Regiones/Paginas/“Veo-bien-aprendo-bien”,-campana-de-salud-visual-impulsada-por-minsalud.aspx>.

- 4] Metro eye, «How to Choose Glasses that Suit You,» 2021. [En línea]. Available: <https://metroeye.biz/how-to-choose-glasses-that-suit-you/>.

- 5] Asociacion Colombiana de oftalmologia, «Uso de gafas de vista en niños,» 2021. [En línea]. Available: <https://acopecolombia.org/uso-de-gafas-en-ninos/>.

- 6] Fundacion Alain Aflelou, «Campana escolar en favor de la salud visual,» 2021. [En línea]. Available: <https://www.afflelou.es/fundacion/proyectos/prevencion-del-fracaso-escolar/> .

- 7] S. N. F. S. S. & K. C. Young, «Eyeglasses frame selection based on oval face shape using convolucional neural network,» *ICIC express letters*, vol. 10, nº 8, pp. 707-715, 2019.

- 8] A. C. a. C. J. M.I Valls Ferrán, «Deteccion precoz de los efectos de refraccion,» *Pediatr Integr*, vol. 17, nº 7, pp. 483-488, 2013.

- 9] Researchgate, «Visión y aprendizaje en la escuela: Guia didactica para la formación de docentes en salud visual infantil,» 2015. [En línea]. Available:

https://www.researchgate.net/publication/279847562_Vision_y_aprendizaje_en_la_Escuela_Guia_didactica_para_la_formacion_de_docentes_en_salud_visual_infantil..

10] Web Bebeymas, «Mi hijo no quiere ponerse las gafas: 11 consejos que le ayudaran a acostumbrarse a ellas,» [En línea]. Available: <https://www.bebesymas.com/consejos/mi-hijo-no-quiere-ponerse-gafas-11-consejos-que-le-ayudaran-a-acostumbrarse-a-ellas>.

11] A. Bezet, «LibGuides: Research process: Datasets,» *libguides*, 2020.

12] J. G. A. M. B. A. U. D. & A. A. R. González, «Metologia para la contruccion de curvas de calidad de estacion,» *Sociedad Española de ciencias*, vol. 18, pp. 303-309, 2004.

13] S. M. a. O. P. P.J. Molina, «JUST-UI: A user interface specification model,» *Springer, Dordrecht*, pp. 63-74, 2002.

14] J. S. G. M. D. C. A. P. B. S. G. M. T. & L. B. b. Perez Carrasco, «Red neuronal convolucional rápida sin fotogramas para reconocimientos de digitos,» *URSI*, pp. 1-4, 2011.

15] Visagismo: harmonia e estética, 6 ed., Sao Pablo: Senac, 2018.

16] Revista Nosotros, «Visagismo: Tipos de rostro».

17] Chibichai, «Guia de tipos de rostros de mujeres,» 2021. [En línea]. Available: <https://chibichai.com/tipos-de-rostros-de-mujeres/>.

18] LIS data solutions, «Deep learning: clasificando imagenes con redes neuronales,» [En línea]. Available: <https://www.lisdatasolutions.com/blog/deep-learning-clasificando-imagenes-con-redes-neuronales/>.

19] A. Lara, Análisis comparativo de arquitecturas de redes neuronales para la clasificación de imagenes, Tesis de maestria Universidad internacional de la Rioja, 2020.

20] Towards data science, «A simple guide to the versions of the inception network,» 2018. [En línea]. Available: <https://towardsdatascience.com/a-simple-guide-to-the-versions-of-the-inception-network-7fc52b863202>.

- 21] H. M and S. M. N, «A review in evaluation metrics for data classification evaluations,» *Int. J. Data min. Knowl. manag. process*, vol. 5, nº 2, pp. 01-11, 2015.
- 22] Towards data science, «The 5 classification evaluation metrics every data scientist must know,» 2019. [En línea]. Available: <https://towardsdatascience.com/the-5-classification-evaluation-metrics-you-must-know-aa97784ff226>.
- 23] U. K. a. J. M. C. S. S. M Viera, «Cohen's kappa coefficient as a performance measure for feature selection,» 2010.
- 24] A. Bazet, «LibGuides: research process: Datasets,» [En línea]. Available: <https://ncu.libguides.com/researchprocess/datasets..>
- 25] J. & L. Zhang, «An interview study on childrens spectacle frame fit. In International Conference on Applied Human Factors and Ergonomics,» pp. 81-88, 2017.
- 26] R. F. ,. S. M. D. A. U. & L. T. Z. Rahmat, «Probabilistic neural network and invariant moments for men face shape classification,» *IOP Conference series: materials science and engineering*, vol. 420, nº 1, p. 012095, Septiembre 2018.
- 27] T. A.-. N. W. & A.-B. B. Alzahrani, «Hybrid feature learning and engineering based approach for face shape classification,» *International conference on intelligent systems and advance computing sciences*, pp. 1-4, 2019.
- 28] H. & V. D. Weerasinghe, «Machine learning approach for hairstyle recommendation,» *International conference on information technology research*, pp. 1-4, 2020.
- 29] A. E. Tio, «Face shape classification using inception,» *ArXiv*, vol. 3, 2019.
- 30] Ascaso Puyuelo, F. J., & Cristóbal Bescós, J. A. (2002). *Sobre lentes, espejuelos, anteojos, gafas o antiparras*. Archivos de la Sociedad Española de Oftalmología, 77(12), 689-691.
- 31] “CiteSeerX — Design Patterns for Research Methods: Iterative Field Research.” <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.535.345> (accessed Mar. 11, 2022).
- 32] A. C. Folgado, «Ciudadanos con la vista de los niños según su edad,» guiainfantil.com, 2017.
- 33] N. Lama, “Face Shape Dataset | Kaggle,” Kaggle, 2020.

- <https://www.kaggle.com/datasets/niten19/face-shape-dataset> (accessed Sep. 23, 2022).
- 34] “Qu’est-ce que l’UI Design et que fait en réalité un UI Designer? | Mobiskill.”
<https://mobiskill.fr/blog/conseils-emploi-tech/quest-ce-que-lui-design-et-que-fait-en-realite-un-ui-designer/> (accessed Sep. 24, 2022).
- 35] J. I. Bagnato, “Convolutional Neural Networks: La Teoría explicada en Español | Aprende Machine Learning,” *Aprendemachinlearning.com*, 2018.
<https://www.aprendemachinlearning.com/como-funcionan-las-convolutional-neural-networks-vision-por-ordenador/> (accessed Sep. 26, 2022).
- 36] “CONCEPTO DE BELLEZA – HISTORIA DE LA SINFONIA.”
<https://www.historiadelasinfonia.es/monografias/las-sinfonias-de-khrennikov/concepto-de-belleza/> (accessed Sep. 29, 2022).